Matches: 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| | :| ||
Db 304 CAIDVKDREC 313

RESULT 3
T09478
ubiquitin thiolesterase (EC 3.1.2.15) - human
N;Alternate names: ubiquitin carboxy-terminal esterase
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: T09478
R;Gray, D.A.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z16684
A;Accession: T09478
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-963 <GRA>
A;Cross-references: EMBL:U20657; NID:g2459396; PID:g2459395
C;Genetics:
A;Gene: Uaph
A;Map position: 3
C;Keywords: proto-oncogene; thiolester hydrolase

Query Match 66.7%; Score 38; DB 2; Length 963;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
:| : |:||
Db 568 CSTSVDGSEC 577

RESULT 4
A24420
notch protein - fruit fly (Drosophila melanogaster)
N;Alternate names: neurogenic repetitive locus protein
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A24420; A24768; S09358; A05267
R;Kidd, S.; Kelley, M.R.; Young, M.W.
Mol. Cell. Biol. 6, 3094-3108, 1986
A;Reference number: A24420; MUID:87064624
A;Accession: A24420
A;Molecule type: DNA
A;Residues: 1-2703 <KID>
A;Cross-references: GB:K03508; NID:g157991; PIDN:AAA28725.1; PID:g157993
R;Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S
Cell 43, 567-581, 1985
A;Reference number: A24768; MUID:86079539
A;Accession: A24768
A;Molecule type: mRNA
A;Residues: 1-48,'I',50-118,'R',120-235,'I',232-256,'N',258-266,'A',268-872,'R',874-958,
A;Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044
R;Tautz, D.
Nucleic Acids Res. 17, 6463-6471, 1989
A;Title: Hypervariability of simple sequences as a general source for polymorphic DNA ma
A;Reference number: S09358; MUID:89385974
A;Accession: S09358
A;Molecule type: DNA
A;Residues: 2505-2551,'QQQC',2553-2576,'E',2578-2604 <TAU>
R;Wharton, K.A.; Yedvobnick, B., Finnerty, V.G.; Artavanis-Tsakonas, S.
Cell 40, 55-62, 1985
A;Title: opa: a novel family of transcribed repeats shared by the Notch locus and other
A;Reference number: A05267; MUID:85099429
A;Accession: A05267
A;Molecule type: DNA
A;Residues: 2504-2576,'E',2578-2611 <WHA2>
C;Genetics:
A;Gene: notch; opa
A;Cross-references: FlyBase:FBan0001647
A;Map position: 8.96-9.36
A;Introns: 54,3-94,[illegible]
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
C;Keywords: differentiation; tandem repeat; transmembrane protein
F;27-43/Domain: transmembrane #status predicted <TMM1>
F;297-328/Domain: EGF homology <EGX1>
F;530-561/Domain: EGF homology <EGF1>
F;568-599/Domain: EGF homology <EGF>
F;988-1019/Domain: EGF homology <EGX2>
F;1064-1095/Domain: EGF homology <EGF3>
F;1187-1218/Domain: EGF homology <EGX3>
F;1746-1762/Domain: transmembrane #status predicted <TMM2>
F;1950-1982/Domain: ankyrin repeat homology <AN1>
F;1983-2015/Domain: ankyrin repeat homology <AN2>
F;1988-2004/Domain: transmembrane #status predicted <TMM3>
F;2017-2049/Domain: ankyrin repeat homology <AN3>
F;2050-2082/Domain: ankyrin repeat homology <AN4>
F;2083-2115/Domain: ankyrin repeat homology <AN5>
F;2538-2568/Region: glutamine-rich
F;2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 66.7%; Score 38; DB 1; Length 2703;
Best Local Similarity 60.0%; Pred. No. 68;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
||| || :|
Db 437 CATGYKGVDC 446

RESULT 5
S20773
Ig heavy chain V region (VH3, 1N48P1) human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 28-Jul-1995
C;Accession: S20773
R;Mortari, F.; Wang, J.; Schroeder, H.W.
submitted to the EMBL Data Library, April 1992
A;Description: Analysis of human cord blood Ig heavy chain VDJ and DJ repertoire.
A;Reference number: S20764
A;Accession: S20773
A;Molecule type: DNA
A;Residues: 1-40 <MOR>
A;Cross-references: EMBL:Z11944; NID:g33878; PIDN:AAA78001.1; PID:g33879
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 64.9%; Score 37; DB 2; Length 40;
Best Local Similarity 60.0%; Pred. No. 2.6;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
| || : || |
Db 8 CTTDLGGAYC 17

RESULT 6
PSLE4E
phospholipase A2 (EC 3.1.1.4) III - broad-banded blue sea krait
N;Alternate names: phosphatidylcholine 2-acylhydrolase
C;Species: Laticauda semifasciata (broad banded blue sea krait, erabu sea snake)
C;Date: 02-Apr-1982 #sequence_revision [illegible] #text_change 28-Feb-1997
C;Accession: A00753
R;Nishida, S.; Kim, H.S.; Tamiya, N.
Biochem. J. 207, 589-594, 1982
A;Title: Amino acid sequences of three phospholipases A I, III and IV from the venom
A;Reference number: A90316; MUID:83153048
A;Accession: A00753
A;Molecule type: protein

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2001, 16:31:02 ; Search time 12.72 Seconds
(without alignments)
59.558 Million cell updates/sec

Title: US-09-018-194-9
Perfect score: 57
Sequence: 1 CAIDIKGAEC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 21641 seqs, 7617552 residues

Total number of hits satisfying chosen parameters: 21641

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	66.7	362	2	I42049	mom 2 protein - Ca
2	38	66.7	381	2	I34254	hypothetical prote
3	38	66.7	963	2	I09478	u/i-puitin thiolest
4	38	66.7	2703	1	A24420	notch protein - fr
5	37	64.9	40	2	S20273	Ig heavy chain V r
6	35	61.4	118	1	PSLT3E	phospholipase A2 (
7	35	61.4	118	1	PSLT4E	phospholipase A2 (
8	35	61.4	118	1	PSLT1E	phospholipase A2 (
9	35	61.4	119	1	PSOXA	phospholipase A2 (
10	35	61.4	404	2	A34832	fat binding protei
11	35	61.4	430	2	I33156	hypothetical prote
12	35	61.4	939	2	T18974	hypothetical prote
13	35	61.4	965	2	I22933	hypothetical prote
14	35	61.4	1048	2	I04172	cadherin-like
15	35	61.4	1124	1	I58388	protein-tyrosine k
16	35	61.4	1125	1	S57846	protein-tyrosine k
17	34.5	60.5	1456	2	I15961	hypothetical prote
18	34	59.6	103	2	S08462	Ig heavy chain V r
19	34	59.6	125	2	A26312	nerve growth facto
20	34	59.6	227	2	B71811	hypothetical prote
21	34	59.6	229	2	I46614	nerve growth facto
22	34	59.6	235	2	S14481	nerve growth facto
23	34	59.6	241	2	JL0097	nerve growth facto
24	34	59.6	242	2	A26312	[illegible] protei
25	34	59.6	245	2	I56570	beta-nerve growth
26	34	59.6	254	2	B55904	multicatalytic end
27	34	59.6	286	1	NGHUBM	nerve growth facto
28	34	59.6	289	2	I42586	alpha-neurotoxyltransl
29	34	59.6	289	2	C42586	alpha-neurotoxyltransl
30	34	59.6	289	2	B42586	alpha-neurotoxyltransl
31	34	59.6	299	2	I75626	cytochrome b-reductase
32	34	59.6	337	1	NCMSMT	beta-1,4-N-acetylgal
33	34	59.6	422	2	S18000	apolipoprotein (H)
34	34	59.6	437	2	B44478	probable poly(ADP-ri
35	34	59.6	441	2	I38072	DNA-directed RNA po
36	34	59.6	478	2	A44478	probable poly(ADP-ri
37	34	59.6	479	2	I34643	ATP-dependent DNA h
38	34	59.6	481	2	B44478	hypothetical prote
39	34	59.6	651	2	I21179	hypothetical prote
40	34	59.6	664	2	I75456	hypothetical prote
41	34	59.6	1040	2	I47801	hypothetical prote
42	33	57.9	84	2	PW1044	ribosomal protein
43	33	57.9	91	2	I44279	cytochrome b-c1 com
44	33	57.9	136	1	CYHUJ6	cytochrome c, juxt
45	33	57.9	139	2	B64565	cytochrome c protein

ALIGNMENTS

RESULT 1
I42049
mom 2 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Dec-1999
C;Accession: I42049
R;Rocheleau, C.E.; Downs, W.D.; Lin, R.; Wittmann, C.; Bei, Y.; Cha, Y.H.; Ali, M.; Priess, J.R.; Mello, C.C.
Cell 90, 707-716, 1997
A;Title: Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos
A;Reference number: A15051; MUID:97433081
A;Accession: I42049
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-362 <Roc>
A;Cross-references: EMBL:AF013952; NID:g2344266; PIDN:AAC47744.1; PID:g2344267
C;Genetics:
A;Gene: mom-2
C;Superfamily: int-1 transforming protein

Query Match 66.7%; Score 38; DB 2; Length 362;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAIDIKGAEC 10
|||:|:|:|| ||
Db 304 CAIDVKDREC 313

RESULT 2
I34254
hypothetical protein F38E1.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jun-2000
C;Accession: I34254
R;Gattuma, S.; Le, L.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid F38E1.
A;Reference number: Z21497
A;Accession: I34254
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-381 <GA>
A;Cross-references: GB:U41006; PIDN:AAA84472.1; PID:g1018911
C;Genetics:
A;Gene: CESP:F38E1.7
A;Map position: 2L, 4.7 +/- 3.7cM, 44.2 +/- 4.0cM
C;Superfamily: int-1 transforming protein

Query Match 66.7%; Score 38; DB 2; Length 381;
Best Local Similarity 66.7%; Pred. No. 12;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2001, 16:31:02 ; Search time 10.04 Seconds
(without alignments)
34.119 Million cell updates/sec

Title: US-09-018-194-9
Perfect score: 57
Sequence: 1 CATDIKGAEC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	66.7	362	1	MOM2_CAEEL	Q10459 caenorhabdi
2	38	66.7	963	1	UBP4_HUMAN	Q13107 homo sapien
3	38	66.7	2703	1	NOTC_DROME	P07207 drosophila
4	35	61.4	118	1	PA21_LATSE	P00611 laticauda s
5	35	61.4	118	1	PA23_LATSE	P00612 laticauda s
6	35	61.4	118	1	PA24_LATSE	P00613 laticauda s
7	35	61.4	119	1	PA21_OXYSC	P00614 oxyuranus s
8	35	61.4	404	1	PRSB_XENLA	Q42586 xenopus lae
9	35	61.4	439	1	PRSA_HUMAN	P17980 homo sapien
10	35	61.4	439	1	PRSA_RAT	Q63569 rattus norv
11	35	61.4	1124	1	TIE2_HUMAN	Q02763 homo sapien
12	35	61.4	1125	1	TIE2_BOVIN	Q06807 bos taurus
13	34	59.6	229	1	NGF_PIG	Q29074 sus scrota
14	34	59.6	231	1	NGF_BOVIN	P13600 bos taurus
15	34	59.6	231	1	NGF_XENLA	P21617 xenopus lae
16	34	59.6	241	1	NGF_CAVPO	P19093 cavia perce
17	34	59.6	241	1	NGF_HUMAN	P01138 homo sapien
18	34	59.6	241	1	NGF_MOUSE	P01139 mus musculu
19	34	59.6	241	1	NGF_RAT	P25427 rattus norv
20	34	59.6	243	1	NGF_CHICK	P05200 gallus gall
21	34	59.6	254	1	PSA7_YEAST	P40303 saccharomyc
22	34	59.6	322	1	ASPG_BACLI	P30363 bacillus li
23	34	59.6	379	1	P2X6_MOUSE	O54803 mus musculu
24	34	59.6	379	1	P2X6_RAT	P51579 rattus norv
25	34	59.6	534	1	UD13_HUMAN	P35503 homo sapien
26	34	59.6	534	1	UD14_HUMAN	P22310 homo sapien
27	34	59.6	534	1	UD15_HUMAN	P35504 homo sapien
28	33	57.9	81	1	YU61_HALNI	O54608 halobacteri
29	33	57.9	106	1	CY2_RHOGL	P00080 rhodopila g
30	33	57.9	129	1	RS11_METJA	P54021 methanococc
31	33	57.9	151	1	Y347_METJA	Q57793 methanococc
32	33	57.9	196	1	IHIJ_ECOLI	Q46948 escherichia
33	33	57.9	208	1	RL13_SCHPO	O74175 schizosacch
34	33	57.9	241	1	NGF_PRANA	P20075 prionys nat
35	33	57.9	279	1	ESR2_MACMU	Q91165 macaca mula
36	33	57.9	365	1	ARNT_HELPO	Q91171 helicobacte
37	33	57.9	530	1	ESR2_CALJA	Q91171 callithrix
38	33	57.9	530	1	ESR2_HUMAN	Q92731 homo sapien
39	33	57.9	824	1	OC27_HUMAN	P30279 homo sapien
40	33	57.9	962	1	UBP4_MOUSE	P35123 mus musculu
41	33	57.9	1068	1	YM54_YEAST	Q04466 saccharomyc
42	33	57.9	1134	1	EIF1_MOUSE	Q60872 mus musculu
43	33	57.9	2437	1	NOTC_BRARE	P46530 brachydan
44	32	56.1	117	1	PA2A_PSEPO	P20258 pseudechis
45	32	56.1	117	1	PA2B_PSEPO	P20259 pseudechis

ALIGNMENTS

RESULT 1
MOM2_CAEEL
ID MOM2_CAEEL STANDARD; PRT; 362 AA.
AC Q10459; Q16146;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MOM-2 PROTEIN PRECURSOR.
GN MOM-2 OR F38E1.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE 97433081; PubMed 9288750;
RA Rocheleau C.E., Downs W.D., Lin R., Wittmann C., Bei Y., Cha Y.-H.,
RA Ali M., Priess J.R., Mello C.C.;
RT "Wnt signaling and an APC-related gene specify endoderm in early C.
RT elegans embryos.";
RL Cell 90:707-716(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Gattung S., Le T.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN INVOLVED IN ENDODERM
CC SPECIFICATION AND CLEAVAGE AXIS DETERMINATION.
CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC ---------- ------ ------- -- -
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- ------ ------- -- -
DR EMBL; AF013952; AAC47749.1; -.
DR EMBL; U41996; AAA83472.1; -.
DR WormPep; F38E1.7; CE04525.
DR InterPro; IPR000970; -.
DR Pfam; PF00110; wnt; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 362 MOM-2 PROTEIN.
FT CARBOHYD 90 90 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 347 362 VESETPEELVLENY... -> SRFLGLGNVCELTEFFPF
FT DLVTVAPFADVHSE (IN REF. 2).
SQ SEQUENCE 362 AA; 42162 MW; E36C13F1F1F1C03 CRC64;

```
  Query Match              66.7%;  Score 38;  DB 1;  Length 962;
  Best Local Similarity    60.0%;  Pred. No. 3.9;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           || |:|  ||
Db     304 CAIDVKDREC 313


RESULT   2
UBP4_HUMAN
ID   UBP4_HUMAN      STANDARD;      PRT;   963 AA.
AC   Q13107; O43452; O44453;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-2008 (Rel. 86, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4 (EC 3.1.2.15) (UBIQUITIN
DE   THIOLESTERASE 4) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 4)
DE   (DEUBIQUITINATING ENZYME 4) (UBIQUITOUS NUCLEAR PROTEIN HOMOLOG).
GN   USP4 OR UNP OR UNPH.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain cortex;
RX   MEDLINE=95303480; PubMed=7784062;
RA   Gray D.A., Inazawa J., Gupta K., Wong A., Ueda R., Takahashi T.;
RT   "Elevated expression of Unph, a proto-oncogene at 3p21.3, in human
RT   lung tumors.";
RL   Oncogene 10:2179-2183(1995).
RN   [2]
RP   REVISIONS.
RA   Gray D.A.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=98124180; PubMed=9464533;
RA   Frederick A., Rolfe M., Chiu M.I.;
RT   "The human UNP locus at 3p21.31 encodes two tissue-selective,
RT   cytoplasmic isoforms with deubiquitinating activity that have reduced
RT   expression in small cell lung carcinoma cell lines.";
RL   Oncogene 16:153-165(1998).
CC   -!- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O =
CC       UBIQUITIN + A THIOL.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; UNPEL (SHOWN HERE) AND UNPES;
CC       ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
CC       FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U20657; AAB72247.1; -.
DR   EMBL; AF017405; AAC27355.1; -.
DR   EMBL; AF017406; AAC27356.1; -.
DR   MEROPS; C19.010; -.
DR   MIM; 603486; -.
DR   InterPro; IPR001394; -.
DR   Pfam; PF00442; UCH-1; 1.
DR   Pfam; PF00443; UCH-2; 1.
DR   PROSITE; PS00972; UCH_2_1; 1.
DR   PROSITE; PS00973; UCH_2_2; 1.
DR   PROSITE; PS50235; UCH_2_3; 1.
KW   Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family;
KW   Proto-oncogene; Alternative splicing.
FT   ACT_SITE    311    311
FT   ACT_SITE    873    873        BY SIMILARITY.
FT   ACT_SITE    881    881        BY SIMILARITY.
FT   VARSPLIC    232    279        KSSTAFSKNFTISFKSSSASEYSSVSASLIANGISISNGMH
FT                                 SSGVSRG -> N (IN ISOFORM UNPES).
FT   MUTAGEN     311    311        C->A: LOSS OF ACTIVITY.
FT   CONFLICT    373    373        S -> R (IN REF. 3).
FT   CONFLICT    744    744        R -> S (IN REF. 3).
SQ   SEQUENCE   963 AA;  108564 MW;  1B62B752E9410CD7 CRC64;


  Query Match              66.7%;  Score 38;  DB 1;  Length 962;
  Best Local Similarity    50.0%;  Pred. No. 9.6;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           |:| : |:||
Db     568 CSTSVDGSEC 577


RESULT   3
NOTC_DROME
ID   NOTC_DROME      STANDARD;      PRT;  2703 AA.
AC   P07207; P04154;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.
GN   N.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86079539; PubMed=3935325;
RA   Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
RT   "Nucleotide sequence from the neurogenic locus Notch implies a gene
RT   product that shares homology with proteins containing EGF-like
RT   repeats.";
RL   Cell 43:567-581(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OREGON R;
RX   MEDLINE=87064624; PubMed=3097517;
RA   Kidd S., Kelley M.R., Young M.W.;
RT   "Sequence of the notch locus of Drosophila melanogaster: relationship
RT   of the encoded protein to mammalian clotting and growth factors.";
RL   Mol. Cell. Biol. 6:3094-3108(1986).
RN   [3]
RP   SEQUENCE OF 2505-2611 FROM N.A.
RX   MEDLINE=85099329; PubMed=2981631;
RA   Wharton K.A., Yedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;
RT   "opa: a novel family of transcribed repeats shared by the Notch locus
RT   and other developmentally regulated loci in D. melanogaster.";
RL   Cell 40:55-62(1985).
RN   [4]
RP   SEQUENCE OF 1-58 FROM N.A.
RX   MEDLINE=87257846; PubMed=3037427;
RA   Kelley M.R., Kidd S., Berg K.L., Young M.W.;
RT   "Restriction of P-element insertions at the Notch locus of Drosophila
RT   melanogaster.";
RL   Mol. Cell. Biol. 7:1545-1548(1987).
RN   [5]
RP   REVIEW.
RA   Harris W.A.;
RT   "Many cell types specified by Notch function.";
RL   Curr. Biol. 1:120-122(1991).
CC   -!- FUNCTION: NOTCH PROTEIN IS ESSENTIAL FOR PROPER DIFFERENTIATION OF
CC       ECTODERM.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM IS
```

```
CC       THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF ONS
CC       DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
CC       NEUROGENIC GENES.
CC   -!- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
CC   -!- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC   -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to licensed@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M16152; AAB59220.1; -.
DR   EMBL; M16153; AAB59220.1; JOINED.
DR   EMBL; M16149; AAB59220.1; JOINED.
DR   EMBL; M16150; AAB59220.1; JOINED.
DR   EMBL; M16151; AAB59220.1; JOINED.
DR   EMBL; K03508; AAA28725.1; -.
DR   EMBL; M13689; AAA28725.1; JOINED.
DR   EMBL; K03507; AAA28725.1; JOINED.
DR   EMBL; M12175; AAA74496.1; -.
DR   EMBL; M19025; AAA28726.1; -.
DR   PIR; A24420; A24420.
DR   PIR; A24768; A24768.
DR   PIR; A05267; A05267.
DR   HSSP; P00740; 1IXA.
DR   FlyBase; FBgn0004647; N.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000800; -.
DR   InterPro; IPR001438; -.
DR   InterPro; IPR001881; -.
DR   InterPro; IPR002110; -.
DR   Pfam; PF00008; EGF; 36.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00066; notch; 3.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 22.
DR   PROSITE; PS00022; EGF_1; 34.
DR   PROSITE; PS01186; EGF_2; 28.
DR   PROSITE; PS01187; EGF_CA; 22.
KW   Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW   Transmembrane; Signal; Glycoprotein.
FT   SIGNAL        1     44       POTENTIAL.
FT   CHAIN        45   2703       NEUROGENIC LOCUS NOTCH PROTEIN.
FT   DOMAIN       45   1745       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1746   1766       POTENTIAL.
FT   DOMAIN     1767   2703       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       58     95       EGF-LIKE 1.
FT   DOMAIN       96    136       EGF-LIKE 2.
FT   DOMAIN      139    176       EGF-LIKE 3.
FT   DOMAIN      177    215       EGF-LIKE 4.
FT   DOMAIN      217    253       EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      255    291       EGF-LIKE 6.
FT   DOMAIN      293    329       EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      331    370       EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      372    408       EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      409    447       EGF-LIKE 10.
FT   DOMAIN      449    486       EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      488    524       EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      526    562       EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      564    600       EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      602    637       EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      639    675       EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      677    713       EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      715    751       EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      753    789       EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      791    827       EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      829    865       EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      867    905       EGF-LIKE 22.
FT   DOMAIN      907    944       EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      946    982       EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      984   1020       EGF-LIKE 25.
FT   DOMAIN     1022   1058       EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1060   1096       EGF-LIKE 27.
FT   DOMAIN     1098   1134       EGF-LIKE 28.
FT   DOMAIN     1136   1181       EGF-LIKE 29.
FT   DOMAIN     1183   1219       EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1221   1257       EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1259   1295       EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     1297   1335       EGF-LIKE 33.
FT   DOMAIN     1337   1373       EGF-LIKE 34.
FT   DOMAIN     1375   1412       EGF-LIKE 35.
FT   DOMAIN     1414   1451       EGF-LIKE 36.
FT   REPEAT     1475   1513       LIN/NOTCH 1.
FT   REPEAT     1514   1553       LIN/NOTCH 2.
FT   REPEAT     1554   1591       LIN/NOTCH 3.
FT   REPEAT     1901   1945       ANK 1.
FT   REPEAT     1959   1979       ANK 2.
FT   REPEAT     1983   2013       ANK 3.
FT   REPEAT     2017   2046       ANK 4.
FT   REPEAT     2050   2079       ANK 5.
FT   REPEAT     2083   2112       ANK 6.
FT   DOMAIN     2538   2568       EGF CLN (OPA-REPEAT).
FT   DISULFID     62     73       BY SIMILARITY.
FT   DISULFID     67     83       BY SIMILARITY.
FT   DISULFID     85     94       BY SIMILARITY.
FT   DISULFID    100    111       BY SIMILARITY.
FT   DISULFID    105    124       BY SIMILARITY.
FT   DISULFID    126    135       BY SIMILARITY.
FT   DISULFID    143    154       BY SIMILARITY.
FT   DISULFID    148    164       BY SIMILARITY.
FT   DISULFID    166    175       BY SIMILARITY.
FT   DISULFID    181    192       BY SIMILARITY.
FT   DISULFID    186    203       BY SIMILARITY.
FT   DISULFID    205    214       BY SIMILARITY.
FT   DISULFID    221    232       BY SIMILARITY.
FT   DISULFID    226    241       BY SIMILARITY.
FT   DISULFID    243    252       BY SIMILARITY.
FT   DISULFID    259    270       BY SIMILARITY.
FT   DISULFID    264    279       BY SIMILARITY.
FT   DISULFID    281    290       BY SIMILARITY.
FT   DISULFID    297    308       BY SIMILARITY.
FT   DISULFID    302    317       BY SIMILARITY.
FT   DISULFID    319    328       BY SIMILARITY.
FT   DISULFID    335    349       BY SIMILARITY.
FT   DISULFID    343    358       BY SIMILARITY.
FT   DISULFID    360    369       BY SIMILARITY.
FT   DISULFID    376    387       BY SIMILARITY.
FT   DISULFID    381    396       BY SIMILARITY.
FT   DISULFID    398    407       BY SIMILARITY.
FT   DISULFID    413    424       BY SIMILARITY.
FT   DISULFID    418    435       BY SIMILARITY.
FT   DISULFID    437    446       BY SIMILARITY.
FT   DISULFID    453    465       BY SIMILARITY.
FT   DISULFID    459    474       BY SIMILARITY.
FT   DISULFID    476    485       BY SIMILARITY.
FT   DISULFID    492    503       BY SIMILARITY.
FT   DISULFID    497    512       BY SIMILARITY.
FT   DISULFID    514    523       BY SIMILARITY.
FT   DISULFID    530    541       BY SIMILARITY.
FT   DISULFID    535    550       BY SIMILARITY.
FT   DISULFID    552    561       BY SIMILARITY.
FT   DISULFID    568    579       BY SIMILARITY.
FT   DISULFID      ?    586       BY SIMILARITY.
FT   DISULFID    590    599       BY SIMILARITY.
FT   DISULFID    606    616       BY SIMILARITY.
FT   DISULFID    611    625       BY SIMILARITY.
FT   DISULFID    627    636       BY SIMILARITY.
FT   DISULFID    643    654       BY SIMILARITY.
```

```
FT   DISULFID    648    663       BY SIMILARITY.
FT   DISULFID    665    674       BY SIMILARITY.
FT   DISULFID    681    692       BY SIMILARITY.

  Query Match             66.7%;  Score 38;  DB 1;  Length 2703;
  Best Local Similarity   60.0%;  Pred. No. 25;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           ||| || :|
Db    437 CATGYKGVDC 446


RESULT   4
PA21_LATSE
ID   PA21_LATSE     STANDARD;      PRT;   118 AA.
AC   P00611;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2 ISOZYME I (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE   2-ACYLHYDROLASE).
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=83153048; PubMed=7165712;
RA   Nishida S., Kim H.S., Tamiya N.;
RT   "Amino acid sequences of three phospholipases A 1, III and IV from
RT   the venom of the sea snake Laticauda semifasciata.";
RL   Biochem. J. 207:589-594(1982).
RN   [2]
RP   REVISIONS TO 70-80.
RX   MEDLINE=89044898; PubMed=3188064;
RA   Takasaki C., Kuramochi H., Shimazu T., Tamiya N.;
RT   "Correction of amino acid sequence of phospholipase A2 I from the
RT   venom of Laticauda semifasciata (Erabu sea snake).";
RL   Toxicon 26:747-749(1988).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   PIR; A00752; PSLFIE.
DR   HSSP; P00603; 1AE7.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
KW   Hydrolase; Lipid degradation; Calcium; Multigene family; Venom.
FT   ACT_SITE     48     48       BY SIMILARITY.
FT   ACT_SITE     64     64       HIGH SPECIFIC ACTIVITY REDUCED 20-FOLD
FT                                BY MODIFICATION OF TRP-64.
FT   ACT_SITE     92     92       BY SIMILARITY.
FT   DISULFID     11     71       BY SIMILARITY.
FT   DISULFID     27    117       BY SIMILARITY.
FT   DISULFID     29     45       BY SIMILARITY.
FT   DISULFID     44     98       BY SIMILARITY.
FT   DISULFID     51     91       BY SIMILARITY.
FT   DISULFID     60     84       BY SIMILARITY.
FT   DISULFID     78     89       BY SIMILARITY.
FT   CA_BIND      49     49       BY SIMILARITY.
SQ   SEQUENCE   118 AA;  13131 MW;  F3BCC44C215B89CD CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 118;
  Best Local Similarity   50.0%;  Pred. No. 5.2;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db     89 CACDLEAAKC 98


RESULT   5
PA23_LATSE
ID   PA23_LATSE     STANDARD;      PRT;   118 AA.
AC   P00612;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2 ISOZYME III (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
DE   ACYLHYDROLASE).
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=83153048; PubMed=7165712;
RA   Nishida S., Kim H.S., Tamiya N.;
RT   "Amino acid sequences of three phospholipases A 1, III and IV from
RT   the venom of the sea snake Laticauda semifasciata.";
RL   Biochem. J. 207:589-594(1982).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   PIR; A00753; PSLIIE.
DR   HSSP; P00608; 1AE7.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
KW   Hydrolase; Lipid degradation; Calcium; Multigene family; Venom.
FT   ACT_SITE     48     48       BY SIMILARITY.
FT   ACT_SITE     92     92       BY SIMILARITY.
FT   DISULFID     11     71       BY SIMILARITY.
FT   DISULFID     27    117       BY SIMILARITY.
FT   DISULFID     29     45       BY SIMILARITY.
FT   DISULFID     44     98       BY SIMILARITY.
FT   DISULFID     51     91       BY SIMILARITY.
FT   DISULFID     60     84       BY SIMILARITY.
FT   DISULFID     78     89       BY SIMILARITY.
FT   CA_BIND      49     49       BY SIMILARITY.
SQ   SEQUENCE   118 AA;  13243 MW;  CCF5671AB8C1560C CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 118;
  Best Local Similarity   50.0%;  Pred. No. 5.2;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db     89 CACDLEAAKC 98


RESULT   6
PA24_LATSE
ID   PA24_LATSE     STANDARD;      PRT;   118 AA.
AC   P00613;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2 ISOZYME IV (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
```

```
DE   ACYLHYDROLASE).
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Fratou sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=83170949; PubMed=7165712;
RA   Nishida S., Kim H.S., Tamiya N.;
RT   "Amino acid sequences of three phospholipases A-1, III and IV from
RT   the venom of the sea snake Laticauda semifasciata.";
RL   Biochem. J. 207:589-594(1982).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   PIR; A90316; PSLT4E.
DR   HSSP; P00608; 1AE7.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
KW   Hydrolase; Lipid degradation; Calcium; Multigene family; Venom.
FT   ACT_SITE     48      48       BY SIMILARITY.
FT   ACT_SITE     92      92       BY SIMILARITY.
FT   DISULFID     11      71       BY SIMILARITY
FT   DISULFID     27     117       BY SIMILARITY.
FT   DISULFID     29      45       BY SIMILARITY.
FT   DISULFID     44      98       BY SIMILARITY.
FT   DISULFID     51      91       BY SIMILARITY.
FT   DISULFID     60      84       BY SIMILARITY.
FT   DISULFID     78      89       BY SIMILARITY.
FT   CA_BIND      49      49       BY SIMILARITY.
SQ   SEQUENCE   118 AA;  13241 MW;  274F9AB18EF0ED3B CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 118;
  Best Local Similarity   50.0%;  Pred. No. 5.2;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db      89 CACDLEAAKC 98


RESULT   7
PA21_OXYSC
ID   PA21_OXYSC     STANDARD;      PRT;   119 AA.
AC   P00614;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2, TAIPOXIN ALPHA CHAIN (EC 3.1.1.4)
DE   (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).
OS   Oxyuranus scutellatus scutellatus (Australian taipan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Oxyuranus.
OX   NCBI_TaxID=8667;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=82261658; PubMed=7049694;
RA   Lind P., Eaker D.;
RT   "Amino acid sequence of the alpha subunit of taipoxin, an extremely
RT   potent presynaptic neurotoxin from the Australian snake taipan
RT   (Oxyuranus s. scutellatus).",
RL   Eur. J. Biochem. 121:441-447(1982).
CC   -!- FUNCTION: TAIPOXIN IS THE MOST POTENT ANIMAL TOXIN KNOWN.
CC       THE ALPHA CHAIN HAS POSSESSES A PHOSPHOLIPASE ACTIVITY.
CC   -!- SUBUNIT: CONTAINS THREE NONCOVALENTLY BOUND CHAINS (ALPHA, BETA,
CC       AND GAMMA), EACH RELATED TO PHOSPHOLIPASE A2.
CC   -!- MISCELLANEOUS: LD(50) IS 0.4 MG/KG BY INTRAVENOUS INJECTION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   PIR; A20754; XJOXA.
DR   HSSP; P00608; 1AE7.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
KW   Hydrolase; Lipid degradation; Calcium; Presynaptic neurotoxin; Venom;
KW   Multigene family.
FT   ACT_SITE     48      48       BY SIMILARITY.
FT   ACT_SITE     92      92       BY SIMILARITY.
FT   DISULFID     11      72       BY SIMILARITY.
FT   DISULFID     27     118       BY SIMILARITY.
FT   DISULFID     29      45       BY SIMILARITY.
FT   DISULFID     44      99       BY SIMILARITY.
FT   DISULFID     51      92       BY SIMILARITY.
FT   DISULFID     60      85       BY SIMILARITY.
FT   DISULFID     78      90       BY SIMILARITY.
FT   CA_BIND      49      49       BY SIMILARITY.
SQ   SEQUENCE   119 AA;  13829 MW;  7FAFD5DD61105C03 CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 119;
  Best Local Similarity   60.0%;  Pred. No. 5.2;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    3;  Gaps    3;

Qy       1 CATDIKGAEC 10
           |  | | ||:
Db      90 CPCDAKAAEC 99


RESULT   8
PRSB_XENLA
ID   PRSB_XENLA     STANDARD;      PRT:   404 AA.
AC   O42586;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   26S PROTEASE REGULATORY SUBUNIT 6A (TAT BINDING PROTEIN 10) (TBP-10).
GN   TBP10.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=98041630; PubMed=9375782;
RA   Nacken W.;
RT   "Members of the AAA gene family are involved in early embryogenesis
RT   of vertebrates.";
RL   Biochim. Biophys. Acta 1354:1-6(1997).
CC   !  FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP DEPENDENT
CC      DEGRADATION OF UBIQUITINATED PROTEINS. THE REGULATORY (19 S) COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC      26S COMPLEX (BY SIMILARITY).
CC   !  SUBUNIT: MAY FORM A HETERODIMER WITH A RELATED FAMILY MEMBER.
CC   !  SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (POTENTIAL).
CC   !  SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC   ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y10460; CAA71486.1; -.
DR   InterPro; IPR001939; -.
DR   Pfam; PF00004; AAA; 1.
DR   PROSITE; PS00674; AAA; 1.
KW   Proteasome; ATP binding; Nuclear protein.
FT   NP_BIND     192    199       ATP (POTENTIAL).
SQ   SEQUENCE   404 AA;  45276 MW;  FC988BBBDCEFC2E3 CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 404;
  Best Local Similarity   50.0%;  Pred. No. 16;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           |  |  ||:|
Db     352 CTDDFNGAQC 361


RESULT   9
PRSA_HUMAN
ID   PRSA_HUMAN      STANDARD;      PRT;   439 AA.
AC   P17980;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   26S PROTEASE REGULATORY SUBUNIT 6A (TAT-BINDING PROTEIN 1) (TBP-1)
DE   (PROTEASOME SUBUNIT P50).
GN   PSMC3 OR TBP1.
OS   Homo Sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93129429; PubMed=8419915;
RA   Ohana B., Moore P.A., Ruben S.M., Southgate C.D., Green M.R.,
RA   Rosen C.A.;
RT   "The type 1 human immunodeficiency virus Tat binding protein is a
RT   transcriptional activator belonging to an additional family of
RT   evolutionarily conserved genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:138-142(1993).
RN   [2]
RP   SEQUENCE OF 46-439 FROM N.A.
RX   MEDLINE=90302011; PubMed=2194290;
RA   Nelbock P., Dillon P.J., Perkins A., Rosen C.A.;
RT   "A cDNA for a protein that interacts with the human immunodeficiency
RT   virus Tat transactivator.";
RL   Science 248:1650-1653(1990).
RN   [3]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=96236387; PubMed=8621709;
RA   Desmaris G.N., Proske R.J., Moomaw C.R., Strong A.A., Song X.,
RA   Hisamatsu H., Tanaka K., Slaughter C.A.;
RT   "Identification, purification, and characterization of a
RT   PA700-dependent activator of the proteasome.";
RL   J. Biol. Chem. 271:3112-3118(1996).
CC   -!- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
CC       DEGRADATION OF UBIQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
CC       COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC       26S COMPLEX (BY SIMILARITY).
CC   -!- FUNCTION: INTERACTS WITH THE HUMAN IMMUNODEFICIENCY VIRUS TAT
CC       TRANSACTIVATOR, SPECIFICALLY SUPPRESSES TAT MEDIATED
CC       TRANSACTIVATION, PROBABLY INVOLVED IN ATP HYDROLYSIS. IT IS
CC       POSSIBLE THAT PROTEINS MSS1 AND TBP-1 COMPETE WITH EACH OTHER TO
CC       REGULATE THE INTERACTION WITH THE TRANSCRIPTIONAL COMPLEX WITH THE
CC       HIV GENE, THEREBY INFLUENCING GENE EXPRESSION IN BOTH DIRECTIONS.
CC   -!- SUBUNIT: MAY FORM A HETERODIMER WITH A RELATED FAMILY MEMBER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is not
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M34079; AAA36666.1; -.
DR   PIR; A34843; A34843.
DR   MIM; 186852; -.
DR   InterPro; IPR001939; -.
DR   Pfam; PF00004; AAA; 1.
DR   PROSITE; PS00674; AAA; 1.
KW   Proteasome; ATP binding; Nuclear protein.
FT   NP_BIND     227    234       ATP (POTENTIAL).
FT   MUTAGEN     233    233       K -> R: LOSS OF FUNCTION.
FT   MUTAGEN     289    289       D -> A: LOSS OF FUNCTION.
SQ   SEQUENCE   439 AA;  49118 MW;  3A9434C5DBBFA69A CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 439;
  Best Local Similarity   50.0%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           |  |  ||:|
Db     387 CTDDFNGAQC 396


RESULT   10
PRSA_RAT
ID   PRSA_RAT        STANDARD;      PRT;   439 AA.
AC   Q63569; P97638;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   26S PROTEASE REGULATORY SUBUNIT 6A (TAT BINDING PROTEIN 1) (TBP-1)
DE   (PROTEASOME SUBUNIT P50/TBP1M ASSOCIATED TAT BINDING 1 PROTEIN B MSS1).
GN   PSMC3 OR TBP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muroidea; Muridae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96188075; PubMed=8607789;
RA   Makino Y., Yoqosawa S., Kamemaki M., Yoshida T., Yamane K.,
RA   Kishimoto T., Moncollin V., Egly J.M., Muramatsu M., Tamura T.;
RT   "Structures of the rat proteasomal ATPases: determination of highly
RT   conserved structural motifs and rules for their spacing.";
RL   Biochem. Biophys. Res. Commun. 220:1049-1054(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE DAWLEY;
RX   MEDLINE=97413049; PubMed=9266764;
RA   Fehrle P., Pfeiffer P.B., Pries F.J., Kierszenbaum A.L.;
RT   "A protein associated with the manchette during rat spermatogenesis is
RT   encoded by a gene of the TBP 1 like subfamily with highly conserved
RT   ATPase and protease domains.";
RL   Mol. Reprod. Dev. 48:77-89(1997).
CC   -!- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
CC       DEGRADATION OF UBIQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
CC       COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
CC       26S COMPLEX (BY SIMILARITY).
CC   -!- SUBUNIT: MAY FORM A HETERODIMER WITH A RELATED FAMILY MEMBER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; D83522; BAA11939.1; -.
DR   EMBL; U77918; AAB70882.1; ALT_INIT.
DR   InterPro; IPR001939; -.
DR   Pfam; PF00004; AAA; 1.
DR   PROSITE; PS00674; AAA; 1.
KW   Proteasome; ATP binding; Nucleat protein.
FT   NP_BIND      227     234       ATP (POTENTIAL).
FT   CONFLICT      84      84       K -> N (IN REF. 2).
FT   CONFLICT     317     317       L -> M (IN REF. 2).
SQ   SEQUENCE   439 AA;  49160 MW;  2928780AEBD19BAE CRC64;


  Query Match             61.4%;  Score 35;  DB 1;  Length 439;
  Best Local Similarity   50.0%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAFDIKGAEC 10
           |  |  ||:|
Db     387 CTDDFNGAQC 396


RESULT  11
TIE2_HUMAN
ID   TIE2 HUMAN      STANDARD;      PRT:  1124 AA.
AC   Q02763;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ANGIOPOIETIN 1 RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE PROTEIN
DE   KINASE RECEPTOR TIE 2) (TYROSINE PROTEIN KINASE RECEPTOR TEK) (P140
DE   TEK) (TUNICA INTERNA ENDOTHELIAL CELL KINASE).
GN   TEK OR TIE2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=93173509; PubMed=8382358;
RA   Kroiler S.F., Bird I.A., Schmeringer J.A., Schooley K.A., Baum P.R.;
RT   "Molecular cloning and characterization of a novel receptor protein
RT   tyrosine kinase from human placenta.";
RL   Oncogene 8:663-670(1993).
RN   [2]
RP   VARIANT VMCM1 TRP-849.
RX   MEDLINE=97134665; PubMed=8980225;
RA   Vikkula M., Boon L.M., Carraway K.L. III, Calvert J.T., Diamonti A.J.,
RA   Goumnerova L., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C.,
RA   Mulliken J.B., Olsen B.R.;
RT   "Vascular dysmorphogenesis caused by an activating mutation in the
RT   receptor tyrosine kinase TIE2.";
RL   Cell 87:1181-1190(1996).
RN   [3]
RP   VARIANTS VMCM1 TRP-849 AND SER-897.
RX   MEDLINE=99299243; PubMed=10369874;
RA   Calvert J.T., Riney T.J., Kontos C.D., Cha E.H., Prieto V.G.,
RA   Chen C.K., Berg J.N., Nevin N.C., Simpson D.A., Pasyk K.A.,
RA   Speer M.C., Peters K.G., Marchuk D.A.;
RT   "Allelic and locus heterogeneity in inherited venous malformations.";
RL   Hum. Mol. Genet. 8:1279-1289(1999).
CC   -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE TRANSMEMBRANE
CC       RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE PREVIOUSLY
CC       MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
CC       ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
CC       PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
CC       FORMATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE  ADP + PHOSPHOPROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN ENDOTHELIAL CELLS
CC       AND THEIR PROGENITORS, THE ANGIOBLASTS. HAS BEEN DIRECTLY FOUND
CC       IN PLACENTA AND LUNG, WITH A LOWER LEVEL IN UMBILICAL VEIN
CC       ENDOTHELIAL CELLS, BRAIN AND KIDNEY.
CC   -!- DISEASE: DEFECTS IN TIE2 ARE A CAUSE OF DOMINANTLY INHERITED
CC       VENOUS MALFORMATIONS (VMCM1), AN ERROR OF VASCULAR MORPHOGENESIS
CC       CHARACTERIZED BY DILATED, SERPIGINOUS CHANNELS.
CC   -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; L06139; AAA61139.1; -.
DR   HSSP; P11362; 1FGI.
DR   MIM; 600221; -.
DR   MIM; 600195; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001777; -.
DR   Pfam; PF00008; EGF; 2
DR   Pfam; PF00041; fn3; 3.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 3.
KW   Receptor; Tyrosine protein kinase; Transferase; Signal; ATP-binding;
KW   Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
KW   Glycoprotein; Phosphorylation; Multigene family; Disease mutation.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1124       ANGIOPOIETIN 1 RECEPTOR.
FT   DOMAIN       19    745       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    746    770       POTENTIAL.
FT   DOMAIN      771   1124       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       44    102       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      210    252       EGF-LIKE 1.
FT   DOMAIN      254    299       EGF-LIKE 2.
FT   DOMAIN      301    341       EGF-LIKE 3.
FT   DOMAIN      370    424       IG-LIKE C2-TYPE DOMAIN.
FT   DOMAIN      444    536       FIBRONECTIN TYPE III.
FT   DOMAIN      541    634       FIBRONECTIN TYPE III.
FT   DOMAIN      638    732       FIBRONECTIN TYPE III.
FT   DOMAIN      824   1096       PROTEIN KINASE.
FT   NP_BIND     830    838       ATP (BY SIMILARITY).
FT   BINDING     855    855       ATP (BY SIMILARITY).
FT   ACT_SITE    964    964       BY SIMILARITY.
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    158    158       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    399    399       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    438    438       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    464    464       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    560    560       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    596    596       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    619    619       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    691    691       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   MOD_RES     992    992       PHOSPHORYLATION (AUTO-).
FT   VARIANT     849    849       R -> W (IN VMCM1; ACTIVATING EFFECT).
FT                                /FTId=VAR_006352.
FT   VARIANT     897    897       Y -> S (IN VMCM1; ACTIVATING EFFECT).
```

```
PT                                  /PTId VAR_008715.
SQ   SEQUENCE   1124 AA;  125810 MW;  65BC05D1BFA4CCEC CRC64;


  Query Match             91.4%;  Score 45;  DB 1;  Length 1124;
  Best Local Similarity   60.0%;  Pred. No. 43;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           ||| || ||
Db     289 CATGWKGLQC 298


RESULT  12
TIE2_BOVIN
ID   TIE2_BOVIN     STANDARD;      PRT;  1125 AA.
AC   Q06807;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ANGIOPOIETIN 1 RECEPTOR PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE   KINASE RECEPTOR TIE-2).
GN   TEK OR TIE2 OR TIE 2.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Endothelial cells;
RX   MEDLINE=94022474; PubMed=8415706;
RA   Sato T.N., Qin Y., Kozak C.A., Audus K.L.;
RT   "Tie-1 and tie-2 define another class of putative receptor tyrosine
RT   kinase genes expressed in early embryonic vascular system.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
CC   -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE TRANSMEMBRANE
CC       RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST
CC       MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
CC       ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
CC       PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
CC       FORMATION.
CC   -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
CC       TYROSINE PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC   -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
CC       ENDOTHELIAL CELLS.
CC   -!- SIMILARITY: TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC
CC       DOMAIN.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X71424; CAA50555.1; -.
DR   PIR; S42691; S42691.
DR   HSSP; P21709; 1FGI.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001777; -.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00041; fn3; 3.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00022; EGF_1; 3.
DR   PROSITE; PS01186; EGF_2; 2.
KW   Receptor; Tyrosine-protein kinase; Transferase; ATP-binding; 3D-structure;
KW   Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
KW   Glycoprotein; Phosphorylation; Multigene family.
FT   SIGNAL        1     18
FT   CHAIN        19   1124       ANGIOPOIETIN 1 RECEPTOR.
FT   DOMAIN       19    745       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    747    771       POTENTIAL.
FT   DOMAIN      772   1124       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       44    102       IG-LIKE C2-TYPE DOMAIN
FT   DOMAIN      210    252       IG-LIKE 1.
FT   DOMAIN      254    299       EGF-LIKE 2.
FT   DOMAIN      301    341       EGF-LIKE 3.
FT   DOMAIN      370    424       IG-LIKE C2-TYPE DOMAIN
FT   DOMAIN      444    539       FIBRONECTIN TYPE III.
FT   DOMAIN      541    635       FIBRONECTIN TYPE III
FT   DOMAIN      639    730       FIBRONECTIN TYPE III
FT   DOMAIN      827   1087       PROTEIN KINASE.
FT   NP_BIND     833    859       ATP (BY SIMILARITY).
FT   BINDING     856    856       ATP (BY SIMILARITY).
FT   ACT_SITE    966    966       BY SIMILARITY.
FT   MOD_RES     993    993       PHOSPHORYLATION (APS  ) (BY SIMILARITY).
FT   CARBOHYD    158    158       N-LINKED (GLCNAC...) (POTENTIAL)
SQ   SEQUENCE   1125 AA;  125927 MW;  0195F1420A80F4D1 CRC64;


  Query Match             91.4%;  Score 45;  DB 1;  Length 1125;
  Best Local Similarity   60.0%;  Pred. No. 43;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           ||| || ||
Db     289 CATGWKGLQC 298


RESULT  13
NGF_PIG
ID   NGF_PIG        STANDARD;      PRT;   229 AA.
AC   Q29074;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   BETA NERVE GROWTH FACTOR PRECURSOR (BETA-NGF) (BETA-NGF).
GN   NGFB.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LARGE WHITE; TISSUE=Blood;
RX   MEDLINE=94311905; PubMed=8016942;
RA   Lahbib-Mansais Y., Mompart F., Yerle M., Gellin J.;
RT   "A new marker (NGFB) on pig chromosome 4, its assignment and
RT   consensus sequence conserved among species.";
RL   Anim. Genet. 25:47-48(1994).
CC   -!- FUNCTION: NERVE GROWTH FACTOR IS IMPORTANT FOR THE DEVELOPMENT AND
CC       MAINTENANCE OF THE SYMPATHETIC AND SENSORY NERVOUS SYSTEMS. IT
CC       STIMULATES DIVISION AND DIFFERENTIATION OF SYMPATHETIC AND
CC       EMBRYONIC SENSORY NEURONS.
CC   -!- SUBUNIT: HOMODIMER, ASSOCIATED BY NON-COVALENT BONDS.
CC   -!- SIMILARITY: BELONGS TO THE NGF-BETA FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   -!- ---------- ----------- ------------------------------------------------------------
DR   EMBL; L41908; AAA21501.1; -.
DR   HSSP; P01139; 1BTG.
DR   InterPro; IPR002072; -.
DR   Pfam; PF00243; NGF; 1.
DR   PROSITE; PS00248; NGF_1; 1.
DR   PROSITE; PS50270; NGF_2; 1.
KW   Growth factor; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1      6       POTENTIAL.
FT   PROPEP        7    109       BY SIMILARITY.
FT   CHAIN       110    229       BETA-NERVE GROWTH FACTOR.
FT   DISULFID    124    189       BY SIMILARITY.
FT   DISULFID    167    217       BY SIMILARITY.
FT   DISULFID    177    219       BY SIMILARITY.
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    102    102       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    154    154       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   229 AA;  25275 MW;  FE8890771CBA3189 CRC64;


  Query Match             59.6%;  Score 34;  DB 1;  Length 229;
  Best Local Similarity   87.5%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           |||||| |
Db     137 ATDIEGKE 144


RESULT  14
NGF_BOVIN
ID   NGF_BOVIN       STANDARD;       PRT;   231 AA.
AC   P13600; Q18969;
DT   01-JAN-1990 (Rel. 13, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   BETA-NERVE GROWTH FACTOR PRECURSOR (BETA-NGF) (FRAGMENT).
GN   NGFB.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=97430845; PubMed=9284944;
RA   Eldridge C., Laurent P., Hayes H., Rodellar C., Leveriel H.,
RA   Zaragoza P.;
RT   "Assignment of the beta-nerve growth factor (NGFB) to bovine
RT   chromosome 3 band q23 by in situ hybridization.";
RL   Cytogenet. Cell Genet. 77:306-307(1997).
RN   [2]
RP   SEQUENCE OF 107-231 FROM N.A.
RX   MEDLINE=86300647; PubMed=2427334;
RA   Meier R., Becker-Andre M., Gotz R., Heumann R., Shaw A., Thoenen H.;
RT   "Molecular cloning of bovine and chick nerve growth factor (NGF):
RT   delineation of conserved and unconserved domains and their
RT   relationship to the biological activity and antigenicity of NGF.";
RL   EMBO J. 5:1489-1493(1986).
CC   -!- FUNCTION: NERVE GROWTH FACTOR IS IMPORTANT FOR THE DEVELOPMENT AND
CC       MAINTENANCE OF THE SYMPATHETIC AND SENSORY NERVOUS SYSTEMS. IT
CC       STIMULATES DIVISION AND DIFFERENTIATION OF SYMPATHETIC AND
CC       EMBRYONIC SENSORY NEURONS.
CC   -!- SUBUNIT: HOMODIMER, ASSOCIATED BY NONCOVALENT FORCES.
CC   -!- SIMILARITY: BELONGS TO THE NGF-BETA FAMILY.
CC   -------- ------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; Y09566; CAA70759.1; -.
DR   EMBL; M26809; AAA30666.1; -.
DR   PIR; A26312; A26312.
DR   HSSP; P01139; 1BTG.
DR   InterPro; IPR002072; -.
DR   Pfam; PF00243; NGF; 1.
DR   PROSITE; PS00248; NGF_1; 1.
DR   PROSITE; PS50270; NGF_2; 1.
KW   Growth factor; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1      8       POTENTIAL.
FT   PROPEP        9    111       BY SIMILARITY.
FT   CHAIN       112    231       BETA NERVE GROWTH FACTOR.
FT   DISULFID    126    191       BY SIMILARITY.
FT   DISULFID    169    219       BY SIMILARITY.
FT   DISULFID    179    221       BY SIMILARITY.
FT   CARBOHYD    154    156       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    118    118       L -> F (IN REF. 2).
FT   CONFLICT    161    161       R -> K (IN REF. 2).
FT   CONFLICT    230    231       AP -> RA (IN REF. 2).
SQ   SEQUENCE   231 AA;  25437 MW;  0160509929 1A4180 CRC64;


  Query Match             59.6%;  Score 34;  DB 1;  Length 231;
  Best Local Similarity   87.5%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           |||||| |
Db     139 ATDIKGKE 146


RESULT  15
NGF_XENLA
ID   NGF_XENLA       STANDARD;       PRT;   231 AA.
AC   P21617;
DT   01-MAY-1991 (Rel. 18, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   NERVE GROWTH FACTOR PRECURSOR (NGF).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91362944; PubMed=1888511;
RA   Carriero F., Campioni M., Cardinali B., Pierandrei-Amaldi P.;
RT   "Structure and expression of the nerve growth factor gene in Xenopus
RT   oocytes and embryos.";
RL   Mol. Reprod. Dev. 29:313-322(1991).
RN   [2]
RP   SEQUENCE OF 170-211 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=91222573; PubMed=2025430;
RA   Hallbrooook F., Ibanez C.F., Persson H.;
RT   "Evolutionary studies of the nerve growth factor family reveal a
RT   novel member abundantly expressed in Xenopus ovary.";
RL   Neuron 6:845-858(1991).
CC   -!- FUNCTION: NERVE GROWTH FACTOR IS IMPORTANT FOR THE DEVELOPMENT AND
CC       MAINTENANCE OF THE SYMPATHETIC AND SENSORY NERVOUS SYSTEMS. IT
CC       STIMULATES DIVISION AND DIFFERENTIATION OF SYMPATHETIC AND
CC       EMBRYONIC SENSORY NEURONS.
CC   -!- SUBUNIT: HOMODIMER, ASSOCIATED BY NONCOVALENT FORCES.
CC   -!- SIMILARITY: BELONGS TO THE NGF-BETA FAMILY.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- ----- ------ ----------------------------------- --- --- ---- --
DR   EMBL; X55710; CAA39249.1; ALT_INIT.
DR   PIR; S14481; S14481.
DR   HSSP; P01139; 1SGF.
DR   InterPro; IPR002072; -.
DR   Pfam; PF00243; NGF; 1.
DR   PRINTS; PR00268; NGF.
DR   PROSITE; PS00248; NGF_1; 1.
DR   PROSITE; PS50270; NGF_2; 1.
KW   Growth factor; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19    114       POTENTIAL.
FT   CHAIN       115    231       NERVE GROWTH FACTOR.
FT   DISULFID    128    193       BY SIMILARITY.
FT   DISULFID    171    221       BY SIMILARITY.
FT   DISULFID    181    223       BY SIMILARITY.
FT   CARBOHYD     63     63       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    107    107       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    158    158       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   231 AA;  26416 MW;  72A04E7D00B85BC5 CRC64;


  Query Match             59.6%;  Score 34;  DB 1;  Length 231;
  Best Local Similarity   87.5%;  Pred. No. 15;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           |||||:|
Db     141 ATDIKGKE 148


Search completed: August 30, 2001, 16:32:00
Job time: 58 sec
```

A;Residues: 1-118 <NIS>
C;Function:
A;Description: catalyzes hydrolysis of 1,2-diacyl-sn-glycero-3-phosphocholine to 1-acyl-
A;Note: the reaction is strongly enhanced when the phospholipid is condensed into a micc
C;Superfamily: phospholipase A2
C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metalloprotein; pros
F;4,67/Binding site: micellar substrate (Glu, Tyr) #status predicted
F;11-71,27-117,29-45,44-98,51-91,60-84,79-89/Disulfide bonds: #status predicted
F;28,30,32,49/Binding site: calcium (Tyr, Gly, Gly, Asp) #status predicted
F;48,92/Active site: His, Asp #status predicted

Query Match 61.4%; Score 35; DB 1; Length 118;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| |:: |:|
Db 89 CACDLEAAKC 98

RESULT 7
PSLT4E
phospholipase A2 (EC 3.1.1.4) IV - broad-banded blue sea krait
N;Alternate names: phosphatidylcholine 2-acylhydrolase
C;Species: Laticauda semifasciata (broad-banded blue sea krait, erabu sea snake)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 28-Feb-1997
C;Accession: A90416; A00753
R;Nishida, S.; Kim, H.S.; Tamiya, N.
Biochem. J. 207, 589-594, 1982
A;Title: Amino acid sequences of three phospholipases A 1, III and IV from the venom of
A;Reference number: A90416; MUID:83153048
A;Accession: A90416
A;Molecule type: protein
A;Residues: 1-118 <NIS>
C;Function:
A;Description: catalyzes hydrolysis of 1,2 diacyl-sn-glycero-3-phosphocholine to 1-acyl-
A;Note: the reaction is strongly enhanced when the phospholipid is condensed into a micc
C;Superfamily: phospholipase A2
C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metalloprotein; pros
F;4,67/Binding site: micellar substrate (Glu, Tyr) #status predicted
F;11-71,27-117,29-45,44-98,51-91,60-84,79-89/Disulfide bonds: #status predicted
F;28,30,32,49/Binding site: calcium (Tyr, Gly, Gly, Asp) #status predicted
F;48,92/Active site: His, Asp #status predicted

Query Match 61.4%; Score 35; DB 1; Length 118;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| |:: |:|
Db 89 CACDLEAAKC 98

RESULT 8
PSLTIE
phospholipase A2 (EC 3.1.1.4) I - broad banded blue sea krait
N;Alternate names: phosphatidylcholine 2-acylhydrolase
C;Species: Laticauda semifasciata (broad banded blue sea krait, erabu sea snake)
C;Date: 02-Apr-1982 #sequence_revision 31-Dec-1989 #text_change 28-Feb-1997
C;Accession: A94325; A00752; A30821
R;Takasaki, C.; Kuramochi, H.; Shimada, T.; Tamiya, N.
Toxicon 26, 747-749, 1988
A;Title: Correction of amino acid sequence of phospholipase A 2 I from the venom of Lati
A;Reference number: A94325; MUID:89044898
A;Accession: A94325
A;Molecule type: protein
A;Residues: 1-118 <TAK>
R;Nishida, S.; Kim, H.S.; Tamiya, N.
Biochem. J. 207, 589-594, 1982
A;Title: Amino acid sequences of three phospholipases A 1, III and IV from the venom of
A;Reference number: A90416; MUID:83153048
A;Contents: annotation; specific activity
A;Note: Trp 64 is important to enzyme activity
C;Function:
A;Description: catalyzes hydrolysis of 1,2 diacyl-sn-glycero-3-phosphocholine to 1-ac
A;Note: the reaction is strongly enhanced when the phospholipid is condensed into a m
C;Superfamily: phospholipase A2
C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metalloprotein; p
F;4,67/Binding site: micellar substrate (Glu, Tyr) #status predicted
F;11-71,27-117,29-45,44-98,51-91,60-84,79-89/Disulfide bonds: #status predicted
F;28,30,32,49/Binding site: calcium (Tyr, Gly, Gly, Asp) #status predicted
F;48,92/Active site: His, Asp #status predicted

Query Match 61.4%; Score 35; DB 1; Length 118;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| |:: |:|
Db 89 CACDLEAAKC 98

RESULT 9
PSOXA
phospholipase A2 (EC 3.1.1.4) taipoxin alpha chain - Australian taipan
N;Alternate names: phosphatidylcholine 2-acylhydrolase
C;Species: Oxyuranus scutellatus scutellatus (Australian taipan)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 28-Feb-1997
C;Accession: A00754
R;Lind, P.; Eaker, D.
Eur. J. Biochem. 124, 441-447, 1982
A;Title: Amino acid sequence of the alpha-subunit of taipoxin, an extremely potent pr
A;Reference number: A00754; MUID:82261658
A;Accession: A00754
A;Molecule type: protein
A;Residues: 1-119 <LIN>
C;Function:
A;Description: catalyzes hydrolysis of 1,2-diacyl-sn-glycero-3-phosphocholine to 1-ac
A;Note: the reaction is strongly enhanced when the phospholipid is condensed into a m
C;Superfamily: phospholipase A2
C;Keywords: calcium; carboxylic ester hydrolase; lipid degradation; metalloprotein; p
F;4,66/Binding site: micellar substrate (Glu, Tyr) #status predicted
F;11-71,27-117,29-45,44-98,51-91,60-84,79-89/Disulfide bonds: #status predicted
F;28,30,32,49/Binding site: calcium (Tyr, Gly, Gly, Asp) #status predicted
F;48,93/Active site: His, Asp #status predicted

Query Match 61.4%; Score 35; DB 1; Length 11-
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
| | | |||
Db 90 CRCDAKAAEC 99

RESULT 10
A34832
fat-binding protein-1 - human
C;Species: Homo sapiens (man)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 19-Jan-2001
C;Accession: A34832
R;Nelbock, P.; Dillon, P.J.; Perkins, A.; Rosen, C.A.
Science 248, 1650-1653, 1990
A;Title: A cDNA for a protein that interacts with the human immunodeficiency virus Tat
A;Reference number: A34832; MUID:90302011
A;Accession: A34832
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-404 <NEL>
A;Cross-references: GB:M34079, NID:g338699; PIDN:AAA36666.1; PID:g338700
C;Superfamily: ATP dependent 26S proteases; P-loop containing nucleoside triphosphate hydrolases; AAA ATPase; AAA+ type ATPase domain
C;Keywords: ATP; nucleotide binding; P-loop

F;165-475/Domain: FtsH/SPC18/VPS48 type ATP-binding domain homologs <VAIP>
F;192-199/Region: nucleotide binding motif A (P loop)

Query Match 61.4%; Score 35; DB 2; Length 404;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAIDIKGAEC 10
Db 352 CIDDFNGAQC 361

RESULT 11
I43155
hypothetical protein F56H1.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C;Accession: I43155
R;Graves, T.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid F56H1.
A;Reference number: Z21293
A;Accession: I43155
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-430 <GRA>
A;Cross-references: EMBL:AF067618; PIDN:AAC19196.1; GSPDB:GN00072; CESP:F56H1.4
A;Experimental source: strain Bristol N2; clone F56H1
C;Genetics:
A;Gene: CESP:F56H1.4
A;Map position: I
A;Introns: 221/3; 438/3; 467/2
C;Superfamily: ATP dependent 26S proteinase; FtsH/SPC18/VPS48 type ATP-binding domain ho

Query Match 61.4%; Score 35; DB 2; Length 429;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAIDIKGAEC 10
Db 378 CIDDFNGAQC 387

RESULT 12
I18074
hypothetical protein C06A1.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: I18074
R;McMurray, A.
submitted to the EMBL Data Library, June 1995
A;Reference number: Z19054
A;Accession: I18074
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-949 <WIL>
A;Cross-references: EMBL:Z49886; PIDN:CAA90054.1; GSPDB:GN00020; CESP:C06A1.4
A;Experimental source: clone C06A1
C;Genetics:
A;Gene: CESP:C06A1.4
A;Map position: I
A;Introns: 52/3; 119/2; 119/3; 282/2; 524/2; 547/3; 644/3; 687/3; 748/3

Query Match 61.4%; Score 35; DB 2; Length 439;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 1; Gaps 0;

Qy 1 CAIDIKGAEC 10
Db 317 CAAVIKGLDC 326

RESULT 13
I22933
hypothetical protein F58G1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: I22933
R;Smye, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19834
A;Accession: I22933
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-965 <WIL>
A;Cross-references: EMBL:Z81556; PIDN:CAB04514.1; GSPDB:GN00[illegible]
A;Experimental source: clone F58G1
C;Genetics:
A;Gene: CESP:F58G1.1
A;Map position: 2
A;Introns: [illegible]

Query Match 61.4%; Score 35; DB 2; Length [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches 6; Conservative 1; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy 1 CAIDIKGAEC 10
Db 344 CAAVIKGLDC 353

RESULT 14
I04172
Ca2+ transporting ATPase (EC 3.6.3.8) - rice
C;Species: oryza sativa (rice)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 07-May-2004
C;Accession: I04172
R;Chen, X.; Chang, M.; Wang, B.; Wu, R.
Plant J. 11, 363-371, 1997
A;Title: Cloning of a Ca(2+) ATPase gene and the [illegible]
A;Reference number: Z15255; MUID:97260072
A;Accession: I04172
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1048 <CHE>
A;Cross-references: EMBL:U82966; NID:g2160711; PIDN:AAB[illegible]
A;Experimental source: cv. IR36
C;Genetics:
A;[illegible]
C;Superfamily: Na+/K+ transporting ATPase alpha chain; ATPase [illegible]
C;Keywords: hydrolase; phosphoprotein
F;39-56/Domain: calcium binding #status predicted <CA1>
F;106-130/Domain: calcium binding #status predicted <CA2>
F;236-256/Domain: calcium binding #status predicted <CA3>
F;[illegible]/Domain: calcium binding #status predicted <CA4>
F;513-696/Domain: ATP binding #status predicted <ATP>
F;610-704/Domain: ATPase nucleotide binding domain [illegible]
F;755-776/Domain: calcium binding #status predicted <CA5>
F;368/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 61.4%; Score 35; DB 2; Length [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches 6; Conservative 1; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy 3 IDIKGAEC 10
Db 197 IDLGKFC 204

RESULT 15
I58388

protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C;Accession: I58388
R;Ziegler, S.F.; Bird, I.A.; Schneringer, J.A.; Schooley, K.A.; Baum, P.R.
Oncogene 8, 663-670, 1993
A;Title: Molecular cloning and characterization of a novel receptor protein tyrosine kin
A;Reference number: I58388; MUID:93173509
A;Accession: I58388
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1124 <RES>
A;Cross-references: GB:L06139; NID:g292823; PIDN:AAA61139.1; PID:g292824
C;Genetics:
A;Gene: GDB:TEK
A;Cross-references: GDB:344185; OMIM:600221
A;Map position: 9p21-9p21
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: protein tyrosine kinase, receptor type tie; EGF homology, fibronectin typ
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1124/Product: short: protein tyrosine kinase, receptor type tek #status predicted <MAT>
F;37-104/Domain: immunoglobulin homology <IM1>
F;145-147/Region: cell attachment (R-G-D) motif
F;211-251/Domain: EGF homology <EG1>
F;255-298/Domain: EGF homology <EG2>
F;302-340/Domain: EGF homology <EG3>
F;364-426/Domain: immunoglobulin homology <IM2>
F;447-527/Domain: fibronectin type III repeat homology <FN3A>
F;542-625/Domain: fibronectin type III repeat homology <FN3B>
F;638-720/Domain: fibronectin type III repeat homology <FN3C>
F;752-772/Domain: transmembrane #status predicted <TMM>
F;822-1099/Domain: protein kinase homology <KIN>
F;840-848/Region: protein kinase ATP-binding motif
F;140,158,399,438,464,560,596,649,691/Binding site: carbohydrate (Asn) (covalent) #statu
F;855,872,964/Active site: Lys, Glu, Asp #status predicted

Query Match 61.4%; Score 35; DB 1; Length 1124;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|||:|| :|
Db 289 CATGWKGLQC 298

Search completed: August 30, 2001, 16:31:44
Job time: 42 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2001, 16:31:02 ; Search time 12.26 Seconds
(without alignments)
16.795 Million cell updates/sec

Title: US-09-018-194-9
Perfect score: 57
Sequence: 1 CAIDIKGAEC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles2.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	66.7	2703	1	US-08-185-432-19	Sequence 19, Appli
2	35	61.4	1124	1	US-08-323-474-2	Sequence 2, Appli
3	35	61.4	1124	5	PCT-US93-06093-2	Sequence 2, Appli
4	34	59.6	9	3	US-08-696-854B-5	Sequence 5, Appli
5	34	59.6	10	3	US-08-696-854B-2	Sequence 2, Appli
6	34	59.6	12	1	US-07-847-369-1	Sequence 1, Appli
7	34	59.6	12	1	US-07-847-369-2	Sequence 2, Appli
8	34	59.6	12	1	US-08-300-044-1	Sequence 1, Appli
9	34	59.6	12	1	US-08-300-044-2	Sequence 2, Appli
10	34	59.6	119	3	US-08-753-642-2	Sequence 2, Appli
11	34	59.6	120	1	US-07-979-630-1	Sequence 1, Appli
12	34	59.6	120	1	US-08-440-049-1	Sequence 1, Appli
13	34	59.6	120	1	US-08-440-049-3	Sequence 3, Appli
14	34	59.6	120	2	US-08-441-513A-1	Sequence 1, Appli
15	34	59.6	120	2	US-08-441-513A-3	Sequence 3, Appli
16	34	59.6	120	3	US-08-970-865-2	Sequence 2, Appli
17	34	59.6	120	3	US-08-970-865-3	Sequence 3, Appli
18	34	59.6	120	3	US-08-581-662-31	Sequence 31, Appl
19	34	59.6	120	4	US-09-363-573-2	Sequence 2, Appli
20	34	59.6	120	4	US-09-363-573-3	Sequence 3, Appli
21	34	59.6	120	5	PCT-US93-11292-1	Sequence 1, Appli
22	34	59.6	120	5	PCT-US95-06918-1	Sequence 1, Appli
23	34	59.6	120	5	PCT-US95-06918-3	Sequence 3, Appli
24	34	59.6	142	1	US-08-451-947-52	Sequence 52, Appl
25	34	59.6	142	2	US-08-424-826A-52	Sequence 52, Appl
26	34	59.6	142	3	US-08-928-694-52	Sequence 52, Appl
27	34	59.6	142	5	PCT-US91-06950-52	Sequence 52, Appl
28	34	59.6	241	1	US-08-266-989B-1	Sequence 1, Appl
29	34	59.6	241	1	US-08-451-947-5	Sequence 5, Appl
30	34	59.6	241	2	US-08-424-826A-5	Sequence 5, Appl
31	34	59.6	241	2	US-08-595-643A-75	Sequence 75, App
32	34	59.6	241	3	US-08-970-865-1	Sequence 1, Appl
33	34	59.6	241	3	US-08-928-694-5	Sequence 5, Appl
34	34	59.6	241	4	US-09-363-573-1	Sequence 1, Appl
35	34	59.6	241	5	PCT-US91-06950-5	Sequence 5, Appl
36	34	59.6	241	5	PCT-US95-05423-4	Sequence 4, Appl
37	34	59.6	289	5	PCT-US92-00282-11	Sequence 11, App
38	34	59.6	289	5	PCT-US92-00282-13	Sequence 13, App
39	34	59.6	289	5	PCT-US92-00282-15	Sequence 15, App
40	34	59.6	379	4	US-09-191-146-32	Sequence 32, App
41	34	59.6	534	5	PCT-US92-00282-4	Sequence 4, Appl
42	33	57.9	131	2	US-08-650-598-3	Sequence 3, Appl
43	33	57.9	485	2	US-08-836-620A-4	Sequence 4, Appli
44	33	57.9	548	4	US-09-139-617-1	Sequence 1, Appli
45	32	56.1	28	3	US-08-482-085B-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-08-185-432-19
; Sequence 19, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Isakonas, Spyridon
; APPLICANT: Bousseau, Isabelle
; APPLICANT: Frederich, Robert J.
; APPLICANT: Su, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: Genetics Reagents, Systems Arrays, and
; TITLE OF INVENTION: Analysis Like, And Related Methods Of Use In Drug
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2703 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-19

Query Match 66.7%, Score 38; DB 1; Length 2703;
Best Local Similarity 66.0%, Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 1; Gaps 0;

```
Qy       1 CAIIIKGAFC 10
           ||| || ||  |
Db     437 CAIGYKGVDC 446


RESULT   2
US-08-323-474-2
; Sequence 2, Application US/08323474
; Patent No. 5447860
;  GENERAL INFORMATION:
;    APPLICANT:  Ziegler, Steven F.
;    TITLE OF INVENTION:  NOVEL TYROSINE KINASE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version # .25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  08/323,474
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/905,600
;      FILING DATE:  26 JUN 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sceso, Kathryn A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2609
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;      TELEX:  756822
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1124 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-323-474-2


  Query Match             61.4%;  Score 35;  DB 1;  Length 1124;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAIIIKGAFC 10
           ||| || ||  |
Db     289 CAIGWKGLQC 298


RESULT   3
PCT/US93/06093
; Sequence 2, Application PCT/US9306093
;  GENERAL INFORMATION:
;    APPLICANT:  Ziegler, Steven F.
;    TITLE OF INVENTION:  NOVEL TYROSINE KINASE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Immunex Corporation
;      STREET:  51 University Street
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/905,600
;      FILING DATE:  26-JUN-1992
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/905,600
;      FILING DATE:  26-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sceso, Kathryn A.
;      REGISTRATION NUMBER:  32,172
;      REFERENCE/DOCKET NUMBER:  2609
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 587-0430
;      TELEFAX:  (206) 233-0644
;      TELEX:  756822
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1124 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US93-06093-2


  Query Match             61.4%;  Score 35;  DB 5;  Length 1124;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAIIIKGAFC 10
           ||| || ||  |
Db     289 CAIGWKGLQC 298


RESULT   4
US-08-696-854B-5
; Sequence 5, Application US/08696854B
; Patent No. 5912078
;  GENERAL INFORMATION:
;    APPLICANT:  SARAGOVI, Uri B.
;    APPLICANT:  LeSAUTEUR, Lynne
;    APPLICANT:  GUELLE, Claudio A.
;    TITLE OF INVENTION:  NERVE GROWTH FACTOR STRUCTURAL ANALOGS
;    TITLE OF INVENTION:  AND THEIR USES
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  KLAUBER & JACKSON
;      STREET:  Continental Plaza, 411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  N.J.
;      COUNTRY:  U.S.A.
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/696,854B
;      FILING DATE:  07-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/CA95/00050
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9402331.4
;      FILING DATE:  07-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JACKSON, David A.
;      REGISTRATION NUMBER:  26,742
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (201) 487-5800
;       TELEFAX:  (201) 343-1684
;       TELEX:  133521
;    INFORMATION FOR SEQ ID NO:  5:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  9 amino acids
;        TYPE:  amino acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  peptide
US-08-696-854B-5


  Query Match             59.6%;  Score 34;  DB 3;  Length 9;
  Best Local Similarity   87.5%;  Pred. No. 1.5e-05,
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           |||||| |
Db       1 ATDIKGKE 8


RESULT    5
US-08-696-854B-2
; Sequence 2, Application US/08696854B
; Patent No. 6017878
;  GENERAL INFORMATION:
;      APPLICANT:  SARAGOVI, Uri H.
;      APPLICANT:  LeSAUTEUR, Lynne
;      APPLICANT:  CUELLO, Claudio A.
;      TITLE OF INVENTION:  NERVE GROWTH FACTOR STRUCTURAL ANALOGS
;      TITLE OF INVENTION:  AND THEIR USES
;      NUMBER OF SEQUENCES:  23
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  KLAUBER & JACKSON
;        STREET:  Continental Plaza, 411 Hackensack Avenue
;        CITY:  Hackensack
;        STATE:  N.J.
;        COUNTRY:  U.S.A.
;        ZIP:  07601
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/08/696,854B
;        FILING DATE:  07-AUG-1996
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  WO PCT/CA95/00059
;        FILING DATE:  07-FEB-1995
;      PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  GB 9402331.4
;        FILING DATE:  07-FEB-1994
;      ATTORNEY/AGENT INFORMATION:
;        NAME:  JACKSON, David A.
;        REGISTRATION NUMBER:  26,742
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  (201) 487-5800
;        TELEFAX:  (201) 343-1684
;        TELEX:  133521
;    INFORMATION FOR SEQ ID NO:  2:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  10 amino acids
;        TYPE:  amino acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;      MOLECULE TYPE:  peptide
;      FEATURE:
;        NAME/KEY:  Modified-site
;        LOCATION:  one-of (1, 10)
;        OTHER INFORMATION:  /note= "Xaa is any uncharged amino acid
;        OTHER INFORMATION:  or hydropathic amino acid"
```

```
US-08-696-854B-2


  Query Match             59.6%;  Score 34;  DB 3;  Length 13;
  Best Local Similarity   85.7%;  Pred. No. 1.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 DIKGAEC 10
           |||| ||
Db       3 DIKGKEC 9


RESULT    6
US-07-847-369-1
; Sequence 1, Application US/07847369
; Patent No. 5349055
;  GENERAL INFORMATION:
;      APPLICANT:  Persson, Hakan B. et al.
;      TITLE OF INVENTION:  Neurotrophic Factors Having Altered
;      TITLE OF INVENTION:  Receptor Binding Specificities
;      NUMBER OF SEQUENCES:  6
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;        STREET:  777 - 84 Saw Mill River Road
;        CITY:  Tarrytown
;        STATE:  New York
;        COUNTRY:  U.S.A.
;        ZIP:  10591-6707
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC-DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER:  US/07/847,369
;        FILING DATE:  19920306
;        CLASSIFICATION:  530
;      ATTORNEY/AGENT INFORMATION:
;        NAME:  Kemple, Gail M./Misrock, S. Leslie
;        REGISTRATION NUMBER:  32,113/18,872
;        REFERENCE/DOCKET NUMBER:  6526-097
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  914-347-7000
;        TELEFAX:  914-347-2113
;    INFORMATION FOR SEQ ID NO:  1:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH:  12 amino acids
;        TYPE:  AMINO ACID
;        STRANDEDNESS:  single
;        TOPOLOGY:  unknown
;      MOLECULE TYPE:  peptide
US-07-847-369-1


  Query Match             59.6%;  Score 34;  DB 1;  Length 12;
  Best Local Similarity   87.5%;  Pred. No. 2.2;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           |||||| |
Db       4 ATDIKGKF 11


RESULT    7
US-07-847-369-2
; Sequence 2, Application US/07847369
; Patent No. 5349055
;  GENERAL INFORMATION:
;      APPLICANT:  Persson, Hakan B. et al.
;      TITLE OF INVENTION:  Neurotrophic Factors Having Altered
;      TITLE OF INVENTION:  Receptor Binding Specificities
;      NUMBER OF SEQUENCES:  6
;      CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;       STREET:  777 Old Saw Mill River Road
;       CITY:  Tarrytown
;       STATE:  New York
;       COUNTRY:  U.S.A.
;       ZIP:  10591-6707
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/847,369
;       FILING DATE:  19920306
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kempler, Gail M./Misrock, S. Leslie
;       REGISTRATION NUMBER:  32,143/18,872
;       REFERENCE/DOCKET NUMBER:  6526-097
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  914 347-7000
;       TELEFAX:  914 347-2113
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-07 847 369 2


  Query Match             59.6%;  Score 34;  DB 1;  Length 12;
  Best Local Similarity   87.5%;  Pred. No. 2.2;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           ||||||| |
Db       4 ATDIKGKE 11


RESULT   8
US 08 300-044-1
; Sequence 1, Application US/08300044
; Patent No. 5705617
;  GENERAL INFORMATION:
;    APPLICANT:  Persson, Hakan B. et al.
;    TITLE OF INVENTION:  Neurotrophic Factors Having Altered
;    TITLE OF INVENTION:  Receptor Binding Specificities
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 Old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10591-6707
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/300,044
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,369
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kempler, Gail M./Misrock, S. Leslie
;      REGISTRATION NUMBER:  32,143/18,872
;      REFERENCE/DOCKET NUMBER:  6526-097
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  914 347-7000
;      TELEFAX:  914 347-2113
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-300-044-1


  Query Match             59.6%;  Score 34;  DB 1;  Length 12;
  Best Local Similarity   87.5%;  Pred. No. 2.2;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           ||||||| |
Db       4 ATDIKGKE 11


RESULT   9
US-08-300-044-2
; Sequence 2, Application US/08300044
; Patent No. 5705617
;  GENERAL INFORMATION:
;    APPLICANT:  Persson, Hakan B. et al.
;    TITLE OF INVENTION:  Neurotrophic Factors Having Altered
;    TITLE OF INVENTION:  Receptor Binding Specificities
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 Old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10591-6707
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/300,044
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,369
;      FILING DATE:  06-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kempler, Gail M./Misrock, S. Leslie
;      REGISTRATION NUMBER:  32,143/18,872
;      REFERENCE/DOCKET NUMBER:  6526-097
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  914 347-7000
;      TELEFAX:  914 347-2113
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-300-044-2


  Query Match             59.6%;  Score 34;  DB 1;  Length 12;
  Best Local Similarity   87.5%;  Pred. No. 2.2;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATDIKGAE 9
           ||||||| |
```

```
Db      4 ATDIKGKE 11


RESULT  10
US-08-753-642-2
; Sequence 2, Application US/08753642
; Patent No. 6063757
;  GENERAL INFORMATION:
;    APPLICANT:  ORSO, RICHARD G.
;    TITLE OF INVENTION:  WOUND TREATMENT METHOD WITH NERVE GROWTH
;    TITLE OF INVENTION:  FACTOR
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS.
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/753,642
;      FILING DATE:  27-NOV-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION.
;      NAME:  MONROY, GLADYS H.
;      REGISTRATION NUMBER:  32,430
;      REFERENCE/DOCKET NUMBER:  32718-20001.00
;    TELECOMMUNICATION INFORMATION.
;      TELEPHONE:  (415) 813-5600
;      TELEFAX:  (415) 494-0792
;      TELEX:  706141 MRSNFOERS SFO
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS.
;      LENGTH:  119 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-753-642-2


  Query Match             59.6%;  Score 34;  DB 3;  Length 119;
  Best Local Similarity   87.5%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AIDIKGAE 9
            |||||| |
Db      29 ATDIKGKE 36


RESULT  11
US-07-979-630-1
; Sequence 1, Application US/07979630
; Patent No. 5488099
;  GENERAL INFORMATION:
;    APPLICANT:  Persson, et al.
;    TITLE OF INVENTION:  Multifunctional Neurotrophic Factors
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Regeneron Pharmaceuticals, Inc.
;      STREET:  777 old Saw Mill River Road
;      CITY:  Tarrytown
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10591
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/979,630
;      FILING DATE:  20-NOV-1992
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,369
;      FILING DATE:  06-MAR-1992
;      NAME:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kempler Ph.D., Gail M.
;      REGISTRATION NUMBER:  32,143
;      REFERENCE/DOCKET NUMBER:  REG 41
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  914-347-7000
;      TELEFAX:  914 347-2113
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  120 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-07-979-630-1


  Query Match             59.6%;  Score 34;  DB 1;  Length 120;
  Best Local Similarity   87.5%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AIDIKGAE 9
            |||||| |
Db      28 ATDIKGKE 35


RESULT  12
US-08-440-049-1
; Sequence 1, Application US/08440049
; Patent No. 5728803
;  GENERAL INFORMATION:
;    APPLICANT:  Urfer, Roman
;    APPLICANT:  Presta, Leonard G.
;    APPLICANT:  Winslow, John W.
;    TITLE OF INVENTION:  PANTROPIC NEUROTROPHIC FACTORS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,049
;      FILING DATE:  12-May-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/253937
;      FILING DATE:  03-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  P0905C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
```

```
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  120 amino acids
;       TYPE:  Amino Acid
;       TOPOLOGY:  Linear
US-08 440 049 1


  Query Match             59.6%;  Score 34;  DB 1;  Length 120;
  Best Local Similarity   87.5%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AIDIKGAE 9
            ||||||| |
Db      28 AIDIKGKE 35


RESULT  14
US-08 440 049-3
; Sequence 3, Application US/08440049
; Patent No. 5728803
;  GENERAL INFORMATION:
;     APPLICANT:  Utter, Roman
;     APPLICANT:  Presta, Leonard G.
;     APPLICANT:  Winslow, John W.
;     TITLE OF INVENTION:  PANTROPIC NEUROTROPHIC FACTORS
;     NUMBER OF SEQUENCES:  7
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  WinPatin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/440,049
;       FILING DATE:  12-May-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/253937
;       FILING DATE:  03-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Torchia, Timothy E.
;       REGISTRATION NUMBER:  36,700
;       REFERENCE/DOCKET NUMBER:  P0905C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-8674
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  120 amino acids
;       TYPE:  Amino Acid
;       TOPOLOGY:  Linear
US-08 440 049 3


  Query Match             59.6%;  Score 34;  DB 1;  Length 120;
  Best Local Similarity   87.5%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AIDIKGAE 9
            ||||||| |
Db      28 AIDIKGKE 35


RESULT  34
US-08 441 513A 2
; Sequence 1, Application US/08441513A
```

```
; Patent No. 5981480
;  GENERAL INFORMATION:
;     APPLICANT:  Utter, Roman
;     APPLICANT:  Presta, Leonard G.
;     APPLICANT:  Winslow, John W.
;     TITLE OF INVENTION:  Pantropic Neurotrophic Factors
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  1 DNA Way
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  WinPatin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  08/253,937
;       FILING DATE:  15 May 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/253937
;       FILING DATE:  03 JUN 1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Torchia, PhD., Timothy E.
;       REGISTRATION NUMBER:  36,700
;       REFERENCE/DOCKET NUMBER:  P0905C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650/225 8674
;       TELEFAX:  650/952 9881
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  120 amino acids
;       TYPE:  Amino Acid
;       TOPOLOGY:  Linear
US-08-441 513A 1


  Query Match             59.6%;  Score 34;  DB 2;  Length 120;
  Best Local Similarity   87.5%;  Pred. No. 23;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AIDIKGAE 9
            ||||||| |
Db      28 AIDIKGKE 35


RESULT  15
US-08 441 513A 3
; Sequence 3, Application US/08441513A
; Patent No. 5981480
;  GENERAL INFORMATION:
;     APPLICANT:  Utter, Roman
;     APPLICANT:  Presta, Leonard G.
;     APPLICANT:  Winslow, John W.
;     TITLE OF INVENTION:  Pantropic Neurotrophic Factors
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  1 DNA Way
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  WinPatin (Genentech)
;     CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/441,513A
; FILING DATE: 15-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/253937
; FILING DATE: 03-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0905C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-441-513A-3

Query Match 59.6%; Score 34; DB 2; Length 120;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAE 9
|||||| |
Db 28 ATDIKGKE 35

Search completed: August 30, 2001, 16:32:47
Job time: 105 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2001, 16:21:02 ; Search time 22.02 Seconds
(without alignments)
60.084 Million cell updates/sec

Title: US-09-018-194-9
Perfect score: 57
Sequence: 1 CATDIKGAEC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	41	71.9	96	5	Q9N7I5	Q9n7i5 leishmania
2	41	71.9	205	5	Q9NDA1	Q9nda1 trypanosoma
3	39	68.4	145	13	Q9I837	Q9i837 laticauda s
4	38	66.7	1035	11	Q9JJG0	Q9jjq0 mus musculu
5	38	66.7	2652	5	Q25253	Q25253 lucilia cup
6	38	66.7	2703	5	Q9W4T8	Q9w4t8 drosophila
7	38	66.7	2704	5	O97458	O97458 drosophila
8	36	63.2	351	11	Q9QZQ5	Q9qzq5 rattus norv
9	35	61.4	145	13	Q9I848	Q9i848 laticauda s
10	35	61.4	145	13	Q9I847	Q9i847 laticauda s
11	35	61.4	145	13	Q9I846	Q9i846 laticauda s
12	35	61.4	145	13	Q9I845	Q9i845 laticauda s
13	35	61.4	145	13	Q9I844	Q9i844 laticauda s
14	35	61.4	145	13	Q9I843	Q9i843 laticauda s
15	35	61.4	145	13	Q9I842	Q9i842 laticauda s
16	35	61.4	194	13	Q9DFE8	Q9dfe8 oncorhynchu
17	35	61.4	319	5	Q96V70	Q96v70 tequla gfei
18	35	61.4	367	14	O72463	O72463 avian ortho
19	35	61.4	367	14	Q65031	Q65031 avian ortho
20	35	61.4	367	14	O12284	O12284 avian reovi
21	35	61.4	367	14	Q91KC3	Q91kc3 avian ortho
22	35	61.4	367	14	Q91KC2	Q91kc2 avian ortho
23	35	61.4	367	14	Q91KC1	Q91kc1 avian ortho
24	35	61.4	367	14	Q91KC0	Q91kc0 avian ortho
25	35	61.4	367	14	Q9DLT2	Q9dlt2 avian ortho
26	35	61.4	367	14	Q9DLP0	Q9dlp0 avian ortho
27	35	61.4	367	14	Q9DL29	Q9dl29 avian ortho
28	35	61.4	367	14	Q9DKH6	Q9dkh6 avian ortho
29	35	61.4	430	5	O76371	O76371 caenorhabdi
30	35	61.4	766	5	Q9V8L3	Q9v8l3 drosophila
31	35	61.4	939	5	Q17685	Q17685 caenorhabdi
32	35	61.4	965	5	O62275	O62275 caenorhabdi
33	35	61.4	975	5	Q9N585	Q9n585 caenorhabdi
34	35	61.4	1048	10	O04938	O04938 oryza sativ
35	35	61.4	1405	5	Q9N694	Q9n694 toxoplasma
36	34.5	60.5	1456	5	Q19127	Q19127 caenorhabdi
37	34	59.6	87	4	Q9P2Z4	Q9p2z4 homo sapien
38	34	59.6	87	6	Q9TIC3	Q9tic3 cervus e.t
39	34	59.6	210	2	Q9Z679	Q9z679 borrelia bj
40	34	59.6	217	6	Q9N183	Q9n183 macaca fas
41	34	59.6	227	2	Q9ZJA5	Q9zja5 helicobacte
42	34	59.6	241	4	Q9UKL8	Q9ukl8 homo sapien
43	34	59.6	241	4	Q9P2Q8	Q9p2q8 homo sapien
44	34	59.6	241	6	Q9N2F1	Q9n2f1 pan troglod
45	34	59.6	241	6	Q9N2F0	Q9n2f0 gorilla gor

ALIGNMENTS

RESULT 1
Q9N7I5
ID Q9N7I5 PRELIMINARY; PRT; 96 AA.
AC Q9N7I5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PROBABLE PROSBELA3 PROTEIN (FRAGMENT).
GN LM28.233.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (?) ?-???-???? to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL390935; CAC00952.1; -.
DR InterPro; IPR001353; -.
DR Pfam; PF00227; proteasome; 1.
FT NON_TER 1 1
FT NON_TER 96 96
SQ SEQUENCE 96 AA; 10774 MW; E9E76B074C9AE991 CRC64;

Query Match 71.9%; Score 41; DB 5; Length 96;
Best Local Similarity 70.0%; Pred. No. 0.8;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
||||: || |
Db 81 CATDLIGAPC 90

RESULT 2
Q9NDA1
ID Q9NDA1 PRELIMINARY; PRT; 205 AA.
AC Q9NDA1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE 20S PROTEASOME BETA 3 SUBUNIT.

GN PSB4.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN 427;
RA Bozdech Z., Huang L., Yao Y., Morton A., Wang C.C.;
RT "TbPSB3 beta 4 subunit of 20S proteasome from T. brucei.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF169653; AAF89685.1; -.
DR InterPro; IPR001353; -.
DR Pfam; PF00227; proteasome; 1.
KW Proteasome.
SQ SEQUENCE 205 AA; 22458 MW; 4CB2092E81E2FFE7 CRC64;

Query Match 71.9%; Score 41; DB 5; Length 205;
Best Local Similarity 70.0%; Prod. No. 1.7;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|||||: || |
Db 122 CATDLIGAPC 131

RESULT 3
Q91837
ID Q91837 PRELIMINARY; PRT; 145 AA.
AC Q91837;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PHOSPHOLIPASE A2 PRECURSOR.
GN PLA2.
OS Laticauda semifasciata (Broad banded blue sea snake) (Erabu sea
OS snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Laticaudinae; Laticauda.
OX NCBI_TaxID=8631;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=VENOM GLAND;
RA Tamiya T., Fujimi T.J.;
RT "Laticauda semifasciata phospholipase A2 cDNA clone LsPLA2cPm09.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tamiya T., Fujimi T.J.;
RT "Laticauda semifasciata phospholipase A2 gene clone LsPLA2Gl 1-2-4th
RT exon.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Tamiya T., Fujimi T.J.;
RT "Laticauda semifasciata phospholipase A2 gene clone LsPLA2Gl2-1 2-4th
RT exon.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB037215; BAB03302.1; -.
DR EMBL; AB037217; BAA99509.1; -.
DR EMBL; AB037216; BAA99509.1; JOINED.
DR EMBL; AB037219; BAA99510.1; -.
DR EMBL; AB037218; BAA99510.1; JOINED.
DR InterPro; IPR001211; -.
DR Pfam; PF00068; phoslip; 1.
DR PRINTS; PR00389; PHPHLIPASEA2.
DR ProDom; PD000303; -; 1.
DR PROSITE; PS00119; PA2_ASP; 1.
DR PROSITE; PS00118; PA2_HIS; 1.
DR SMART; SM00085; PA2c; 1.
KW Signal.
FT SIGNAL 1 27 POTENTIAL.

SQ SEQUENCE 145 AA; 16099 MW; A6245E7C0D0ED38F CRC64;

Query Match 68.4%; Score 39; DB 13; Length 145;
Best Local Similarity 60.0%; Prod. No. 3;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| |:| |:|
Db 116 CAPCFAAFC 125

RESULT 4
Q9JJG0
ID Q9JJG0 PRELIMINARY; PRT; 1035 AA.
AC Q9JJG0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE BRAIN cDNA, CLONE MNCB-4522, SIMILAR TO AF220015, TAX2 (HOMO SAPIENS)
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL;
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sudano S.,
RA Hashimoto K.;
RT "Isolation of full length cDNA clones from mouse brain cDNA library
RT made by oligo-capping method.";
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB041546; BAA95041.1; -.
DR InterPro; IPR002965; -.
DR PRINTS; PR01217; PRICHEXTENSN.
SQ SEQUENCE 1035 AA; 112715 MW; 0698015282F94C78 CRC64;

Query Match 66.7%; Score 38; DB 11; Length 1035;
Best Local Similarity 60.0%; Prod. No. 32;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|: | || ||
Db 134 CSADSKGQEC 143

RESULT 5
Q25253
ID Q25253 PRELIMINARY; PRT; 2653 AA.
AC Q25253;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE NOTCH HOMOLOG SCALLOPED WINGS (SCL).
GN SCL.
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Lucilia.
OX NCBI_TaxID=7375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SS SEEKING;
RX MEDLINE=96400928; PubMed=8807304;
RA Davies A.G., Game A.Y., Chen S., Williams T.J., Goodall S., Yen J.J.,
RA McKenzie J.A., Batterham P.;
RT "Scalloped wings is the Lucilia cuprina Notch homologue and a
RT candidate for the modifier of fitness and asymmetry of diazinon
RT resistance.";
RL Genetics 143:1421-1437(1996).
RN [2]
RP SEQUENCE OF 39-265 FROM N.A.

```
RC   STRAIN=SS SEEKING;
RA   Chen Z., Newsome T., McKenzie J.A., Batterham P.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 39-265 FROM N.A.
RC   STRAIN=SS SEEKING;
RA   Chen Z., McKenzie J.A., Batterham P.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D58977; AAC36151.1; -.
DR   EMBL; AF042672; AAC36152.1; -.
DR   EMBL; AF042670; AAC36152.1; JOINED.
DR   EMBL; AF042671; AAC36152.1; JOINED.
DR   EMBL; AF042673; AAC36153.1; -.
DR   HSSP; P00740; 1EDM.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000742; -.
DR   InterPro; IPR000800; -.
DR   InterPro; IPR001438; -.
DR   InterPro; IPR001881; -.
DR   InterPro; IPR002110; -.
DR   Pfam; PF00008; EGF; 36.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00066; notch; 3.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 22.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_34.
DR   PROSITE; PS01186; EGF_2; 28.
DR   PROSITE; PS01187; EGF_CA; 21.
DR   SMART; SM00248; ANK; 1.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ   SEQUENCE   2653 AA;  285928 MW;  6AF2A058FFFC6329 CRC64;
```

```
  Query Match             66.7%;  Score 38;  DB 5;  Length 2653;
  Best Local Similarity   60.0%;  Pred. No. 80;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           ||| || :|
Db     418 CATGYKGVDC 427
```

```
RESULT   6
Q9W4T8
ID   Q9W4T8       PRELIMINARY;      PRT:  2703 AA.
AC   Q9W4T8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   N GENE PRODUCT.
GN   N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE 20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixen K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003426; AAF45848.2; -.
DR   HSSP; P00740; 1IXA.
DR   FlyBase; FBgn0034647; N.
DR   InterPro; IPR000152; -.
DR   InterPro; IPR000561; -.
DR   InterPro; IPR000742; -.
DR   InterPro; IPR000800; -.
DR   InterPro; IPR001438; -.
DR   InterPro; IPR001881; -.
DR   InterPro; IPR002049; -.
DR   InterPro; IPR002110; -.
DR   Pfam; PF00008; EGF; 36.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00066; notch; 3.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00011; EGFLAMININ.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 22.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_34.
DR   PROSITE; PS01186; EGF_2; 28.
DR   PROSITE; PS01187; EGF_CA; 21.
KW   Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ   SEQUENCE   2703 AA;  288851 MW;  0EAE23F426FEC07B CRC64;
```

```
  Query Match             66.7%;  Score 38;  DB 5;  Length 2703;
  Best Local Similarity   60.0%;  Pred. No. 81;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           ||| || :|
Db     437 CATGYKGVDC 446
```

```
RESULT   7
O97458
ID   O97458       PRELIMINARY;      PRT:  2704 AA.
AC   O97458;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   EG:163A10.2 PROTEIN.
GN   N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
OS   Drosophila melanogaster (Fruit fly).
```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Borkova D., Minana B., Katatos F.C.;
RT "Sequencing the distal X chromosome of Drosophila melanogaster.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benos P.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL035439; CAB37611.1; -.
DR EMBL; AL035436; CAB37611.1; JOINED.
DR EMBL; AL035436; CAB37610.1; -.
DR EMBL; AL035439; CAB37610.1; JOINED.
DR HSSP; P00740; 1EDM.
DR FlyBase; FBgn0004647; N.
DR InterPro; IPR000152; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000742; -.
DR InterPro; IPR000800; -.
DR InterPro; IPR001438; -.
DR InterPro; IPR001881; -.
DR InterPro; IPR002049; -.
DR InterPro; IPR002110; -.
DR Pfam; PF00008; EGF; 36.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PR00010; EGFBLOOD.
DR PRINTS; PR00011; EGFLAMININ.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; UNKNOWN_34.
DR PROSITE; PS01186; EGF_2; 28.
DR PROSITE; PS01187; EGF_CA; 21.
DR SMART; SM00179; EGF_CA; 1.
KW Calcium binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2704 AA; 288894 MW; 554298A25C9EBA4E CRC64;

Query Match 66.7%; Score 38; DB 5; Length 2704;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CATDIKGAEC 10
|||:|| :|
Db 437 CATGYKGVDC 446

RESULT 8
Q9QZQ5
ID Q9QZQ5 PRELIMINARY; PRT; 451 AA.
AC Q9QZQ5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE NOV PROTEIN.
GN NOV.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN SPRAGUE-DAWLEY;
RX MEDLINE=20035752; PubMed=10570975;
RA Lin C., Liu X.J., Crowe P.D., Kelner G.S., Fan J., Barry G., Manu F.,
RA Lina N., De Souza E.B., Maki R.A.;
RT "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
RT that induces protein tyrosine phosphorylation.";
RL Gene 238:471-478(1999).
DR EMBL; AF171936; AAD43371.1; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000867; -.
DR InterPro; IPR000884; -.
DR InterPro; IPR001007; -.
DR Pfam; PF00007; cys_knot; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00219; IGFBP; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS01208; VWFC; UNKNOWN 1.
DR SMART; SM00041; CT; 1.
SQ SEQUENCE 451 AA; 48709 MW; 02E29707DE7C1BEF CRC64;

Query Match 64.2%; Score 36; DB 11; Length 451;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAEC 10
|||:|| :|
Db 250 ATDMKGKKC 258

RESULT 9
Q91848
ID Q91848 PRELIMINARY; PRT; 145 AA.
AC Q91848;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PHOSPHOLIPASE A2.
GN PLA2.
OS Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Laticaudinae; Laticauda.
OX NCBI_TaxID=8631;
RN [1]
RP SEQUENCE FROM N.A.
RA Tamiya T., Fujimi T.J.;
RT "Laticauda semifasciata phospholipase A2 gene cloning [sPLA2G1] 1 / 4 th
RT exon.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB037221; BAA99511.1; -.
DR EMBL; AB037220; BAA99511.1; JOINED.
DR InterPro; IPR001211; -.
DR Pfam; PF00068; phoslip; 1.
DR PRINTS; PR00389; PHPHLIPASEA2.
DR PROSITE; PS00119; PA2_ASP; 1.
DR PROSITE; PS00118; PA2_HIS; 1.
DR SMART; SM00085; PA2c; 1.
SQ SEQUENCE 145 AA; 15900 MW; 1E841E6F7B0C4481 CRC64;

Query Match 61.4%; Score 35; DB 15; Length 145;
Best Local Similarity 50.0%; Pred. No. 196;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|| |:: :|:|
Db 116 CACDLEAAKC 125

RESULT 10
Q91847
ID Q91847 PRELIMINARY; PRT; 145 AA.
AC Q91847;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR
GN   PLA2.
OS   Laticauda semifasciata (Broad banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA clone LsPLA2cL037.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037409; BAB03296.1; -.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL       1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16062 MW;  D3BC01790A0F7885 CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAFC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


RESULT  11
Q91846
ID   Q91846      PRELIMINARY;      PRT;   145 AA.
AC   Q91846;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR.
GN   PLA2.
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA clone LsPLA2cL038.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037410; BAB03297.1; -.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL       1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16059 MW;  9277AC82B50F7887 CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAFC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


RESULT  12
Q91845
ID   Q91845      PRELIMINARY;      PRT;   145 AA.
AC   Q91845;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR.
GN   PLA2.
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA clone LsPLA2cL036.";
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037411; BAB03298.1; -.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL       1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16089 MW;  9277B0F2B07F6D87 CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAFC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


RESULT  13
Q91844
ID   Q91844      PRELIMINARY;      PRT;   145 AA.
AC   Q91844;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR.
GN   PLA2.
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA clone LsPLA2cP110.";
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB037412; BAB03299.1; -.
```

```
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL        1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16073 MW;  49D4B439160F7895 CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


RESULT  14
Q91843
ID   Q91843      PRELIMINARY;      PRT;   145 AA.
AC   Q91843;
DT   01 OCT-2000 (TrEMBLrel. 15, Created)
DT   01 OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01 MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR.
GN   PLA2.
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA clone LSFA2cPm05.";
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB047413; BAB03300.1; -.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL        1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16073 MW;  926F0F8886IDE887 CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


RESULT  15
Q91842
ID   Q91842      PRELIMINARY;      PRT;   145 AA.
AC   Q91842;
DT   01 OCT 2000 (TrEMBLrel. 15, Created)
DT   01 OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   PHOSPHOLIPASE A2 PRECURSOR.
GN   PLA2.
OS   Laticauda semifasciata (Broad-banded blue sea snake) (Erabu sea
OS   snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; [illegible]
OC   Elapidae; Laticaudinae; Laticauda.
OX   NCBI_TaxID=8631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=VENOM GLAND;
RA   Tamiya T., Fujimi T.J.;
RT   "Laticauda semifasciata phospholipase A2 cDNA [illegible]";
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB047414; BAB03301.1; -.
DR   InterPro; IPR001211; -.
DR   Pfam; PF00068; phoslip; 1.
DR   PRINTS; PR00389; PHPHLIPASEA2.
DR   ProDom; PD000303; -; 1.
DR   PROSITE; PS00119; PA2_ASP; 1.
DR   PROSITE; PS00118; PA2_HIS; 1.
DR   SMART; SM00085; PA2c; 1.
KW   Signal.
FT   SIGNAL        1     27       POTENTIAL.
SQ   SEQUENCE   145 AA;  16085 MW;  66A2CED87AEGED6C CRC64;


  Query Match             61.4%;  Score 35;  DB 13;  Length 145;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           || |:: |:|
Db     116 CACDLQAAKC 125


Search completed: August 30, 2001, 16:42:29
Job time: 87 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 30, 2001, 16:39:57 ; Search time 18.14 Seconds
(without alignments)
33.420 Million cell updates/sec

Title: US-09-018-194-9
Perfect score: 57
Sequence: 1 CATDIKGAEC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623488 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]
7: [illegible]
8: [illegible]
9: [illegible]
10: [illegible]
11: [illegible]
12: [illegible]
13: [illegible]
14: [illegible]
15: [illegible]
16: [illegible]
17: [illegible]
18: [illegible]
19: [illegible]
20: [illegible]
21: [illegible]
22: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	100.0	10	20	AAY39234	Cyclic peptide use
2	40	70.2	811	21	AAB51727	Human secreted pro
3	35	61.4	794	22	AAB31165	Amino acid sequenc
4	35	61.4	977	22	AAB31164	Amino acid sequenc
5	35	61.4	1124	15	AAR45440	Human orphan recep
6	35	61.4	1124	16	AAR73953	Human TEK tyrosine
7	35	61.4	1124	20	AAY30318	Amino acid sequenc
8	34.5	60.5	1296	20	AAW88448	Chemerin/tazarotin
9	34	59.6	9	16	AAR83764	NGF derived struct
10	34	59.6	10	16	AAR83761	NGF derived struct
11	34	59.6	11	10	AAP90624	Peptide to induce
12	34	59.6	12	14	AAR41747	Glycosylated NGF p
13	34	59.6	12	14	AAR41748	Glycosylated NGF p
14	34	59.6	13	11	AAR07158	Nerve growth facto
15	34	59.6	14	11	AAR07161	Synthetic Nerve gr
16	34	59.6	24	12	AAR13889	(28-47)hNGF. Synt
17	34	59.6	27	12	AAR13893	(9-35)hNGF. Synth
18	34	59.6	27	12	AAR13894	(37-62)hNGF. Synt
19	34	59.6	78	12	AAR15015	Beta lactamase sti
20	34	59.6	118	10	AAP91034	Human nerve growth
21	34	59.6	118	13	AAR21874	Chimeric recombina
22	34	59.6	118	13	AAR29493	NGF, mouse. Mus m
23	34	59.6	119	5	AAP40040	Sequence encoded b
24	34	59.6	119	10	AAP90133	Human nerve growth
25	34	59.6	119	12	AAR13864	Human NGF B-chain
26	34	59.6	119	12	AAR13872	Variant human beta
27	34	59.6	119	16	AAR77420	Nerve growth [illegible]
28	34	59.6	119	21	AAB03347	Human beta nerve g
29	34	59.6	120	13	AAR21863	Chimeric beta-NGF
30	34	59.6	120	13	AAR21864	Chimeric beta-NGF
31	34	59.6	120	13	AAR21867	Chimeric beta-NGF
32	34	59.6	120	13	AAR21868	Chimeric beta-NGF
33	34	59.6	120	13	AAR21869	Chimeric beta-NGF
34	34	59.6	120	13	AAR21870	Chimeric beta-NGF
35	34	59.6	120	13	AAR21871	Chimeric beta-NGF
36	34	59.6	120	13	AAR21875	Chimeric beta-NGF
37	34	59.6	120	15	AAR54084	Nerve growth fact
38	34	59.6	120	17	AAR90743	Rabbit pro beta-N
39	34	59.6	120	17	AAR90753	Rabbit pro beta-N
40	34	59.6	120	19	AAW48867	Mouse beta-NGF pro
41	34	59.6	120	20	AAW81134	Nerve growth fact
42	34	59.6	120	20	AAW81135	Nerve growth fact
43	34	59.6	120	20	AAW81136	Nerve growth fact
44	34	59.6	120	20	AAW81137	Nerve growth fact
45	34	59.6	120	20	AAW81138	Nerve growth fact

ALIGNMENTS

RESULT 1
AAY39234
ID AAY39234 standard; peptide; 10 AA.
XX
AC AAY39234;
XX
DT [illegible]
XX
DE Cyclic peptide used to inhibit p75NTR mediated apoptosis.
XX
KW Cyclic peptide; p75NTR; p75 neurotrophin receptor; nerve growth fact
KW CNT mediated; keratinocyte; apoptosis; Bcl-2; beta-amyloid;
KW Alzheimer's disease; pseudo-ligand; hair growth; hair colour;
KW skin colour; alopecia areata; male pattern baldness.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note "N-terminus is attached via a peptide bond to
FT the C-terminus amino acid at position 10"
FT Modified-site 10
FT /note "C-terminus is attached via a peptide bond to
FT the N-terminus amino acid at position 1"
XX
PN WO9939728-A2.
XX
PD 12-AUG-1999.
XX
PF 03-FEB-1999; 99WO-US002362.
XX
PR 04-FEB-1998; 98US-0018194.
XX
PA (UYBI-) UNIV BOSTON.

XX
PT Eller M, Gilchrest BA, Yaar M;
XX
DR WPI; 1999-539950/45.
XX
PT Controlling or manipulating melanocyte and keratinocyte cell death,
PT useful for treating, e.g. alopecia areata
XX
PS Claim 4; Page 40; 67pp; English.
XX
CC This sequence is a cyclic peptide which competes with and competitively
CC inhibits beta-amyloid binding to the p75 neurotrophin receptor (p75NTR).
CC P75NTR is a low affinity nerve growth factor (NGF) receptor which is
CC expressed by melanocytes and keratinocytes of the basal epidermis.
CC Apoptosis can be inhibited by p75NTR via the upregulation of the Bcl-2
CC protein. If the receptor is occupied by appropriate ligands e.g.
CC neurotrophins, apoptosis is inhibited. Other examples of appropriate
CC ligands include this cyclic peptide. This peptide is based on the
CC sequence of the binding fragment of beta amyloid, which binds to
CC p75NTR in Alzheimer's disease. This cyclic peptide can be used in
CC methods to control or manipulate keratinocyte or melanocyte cell death.
CC The methods involve using this peptide or peptides AA[illegible] to
CC bind as a pseudo-ligand to the p75 neurotrophin receptor, inhibiting
CC apoptosis. The new method can induce or maintain hair growth, hair colour
CC or skin colour. Inducing or maintaining hair growth is useful for
CC treating alopecia areata or male pattern baldness in vertebrates.
XX
SQ Sequence 10 AA;

Query Match 100.0%; Score 57; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAIDIRGAEC 10
||||||||||
Db 1 cardikqaec 10

RESULT 2
AAB51727
ID AAB51727 standard; Protein; 311 AA.
XX
AC AAB51727;
XX
DT 16-FEB-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 4.
XX
KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW opthalmological; vulnerary; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorders; cancer; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; nervous system disorder;
KW Alzheimer's disease; ocular disorder; wound healing; skin ailment.
XX
OS Homo sapiens.
XX
PN WO200061625-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US08981.
XX
PR 09-APR-1999; 99US-0128701.
PR 20-JAN-2000; 2000US-0177166.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-619226/59.
XX
PT New nucleic acid molecules encoding 48 human secreted proteins, for
PT diagnosing, preventing, treating or ameliorating medical conditions, and
PT used as food additives or preservatives
XX
PS Disclosure; Page 446; 509pp; English.
XX
CC Polynucleotide sequences AA[illegible] AA[illegible] represent cDNA encoding
CC human secreted proteins AAB51724 - AAB51777. Sequence AAB[illegible]
CC AAB51825 represent alternative polypeptides encoded by the genes, and
CC amino acid sequences to which they are homologous. The genes and proteins
CC have activities dependent on the tissues and cells in which they are
CC expressed. Examples of these activities include immunosuppressive,
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
CC virucide; fungicide; opthalmological; and vulnerary. The secreted
CC proteins, polynucleotides, antagonists and agonists may be useful in
CC treating, preventing and/or diagnosing diseases and disorders such as
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral infarction,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin, mucous and
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative, to
CC increase or decrease storage capabilities, for certain rapid protein
CC carbohydrate, vitamins, minerals, cofactors and other nutritional
CC components. oligonucleotide AA[illegible] AA[illegible] and peptide AA[illegible] are
CC used in the isolation and characterisation of the proteins and
CC polynucleotides of the invention.
XX
SQ Sequence 311 AA;

Query Match 70.2%; Score 40; DB 21; Length [illegible];
Best Local Similarity 90.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAIDIRGAEC 10
|:|||||:|
Db 252 cstdvrqahc 261

RESULT 3
AAB31165
ID AAB31165 standard; protein; 704 AA.
XX
AC AAB31165;
XX
DT 02-APR-2001 (first entry)
XX
DE Amino acid sequence of a Tek/Fc fusion protein.
XX
KW Fusion protein; receptor tyrosine kinase; Tek; Tie2; [illegible];
KW immunoglobulin G1; IgG1; angiogenesis; tumour; [illegible]; vasculitis; [illegible];
KW inflammatory disease; arthritis; rheumatism; psoriasis; [illegible];
KW diabetic retinopathy; retinopathy; neovascular glaucoma; [illegible];
KW retrolental fibroplasia; rubeosis; uveitis; macular degeneration;
KW graft neovascularisation; cancer; metastatic carcinoma; [illegible];
KW wound granulation.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /note= "signal peptide"
FT Protein 19..474
FT /note= "Tek extracellular domain fragment"

```
FT   Protein          473..704
FT                    /note= "Fc portion"
XX
PN   WO200075323-A1.
XX
PD   14-DEC-2000.
XX
PF   07-JUN-2000; 2000WO-US15706.
XX
PR   07-JUN-1999;   99US-0137889.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Cerretti DP,  Borges LG,  Fanslow WC;
XX
DR   WPI; 2001-112149/12.
XX
PT   New tek polypeptides antagonist having a fragment of the Tek
PT   extracellular domain, useful for treating mammals with a disease
PT   mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
PT   inflammatory diseases -
XX
PS   Claim 17; Page 47-49; 43pp; English.
XX
CC   The present sequence represents a fusion protein comprising the
CC   extracellular domain of the human receptor tyrosine kinase Tek and the
CC   Fc portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all
CC   or part of the region containing fibronectin type III (FNIII) motifs, and
CC   retains the ability to bind at least one Tek ligand. The fusion
CC   polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The
CC   polypeptide or soluble Tek multimer, antibodies or antibody fragments are
CC   useful for treating a mammal having a disease or condition mediated by
CC   angiogenesis, e.g. a solid tumour a condition or disease characterised by
CC   ocular neovascularisation. In particular, the Tek antagonists are useful
CC   for treating or preventing inflammatory diseases (e.g. arthritis,
CC   rheumatism or psoriasis), certain eye disorders (e.g. diabetic
CC   retinopathy, retinopathy of prematurity, neovascular glaucoma,
CC   retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
CC   degeneration or graft neovascularisation), cancer (e.g. metastatic
CC   sarcomas or carcinomas), or wound granulation.
XX
SQ   Sequence   704 AA;
```

```
  Query Match             61.4%;  Score 35;  DB 22;  Length 704;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           |||  || :|
Db     289 catqwkqlqc 298
```

```
RESULT   4
AAB31164
ID   AAB31164 standard; protein; 977 AA.
XX
AC   AAB31164;
XX
DT   02-APR-2001  (first entry)
XX
DE   Amino acid sequence of a Tek/Fc fusion protein.
XX
KW   Fusion protein; receptor tyrosine kinase; Tek; Fc portion;
KW   immunoglobulin G1; IgG1; angiogenesis; tumour; ocular neovascularisation;
KW   inflammatory disease; arthritis,; rheumatism; psoriasis; eye disorder;
KW   diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;
KW   retrolental fibroplasia; rubeosis; uveitis; macular degeneration;
KW   graft neovascularisation; cancer; metastatic sarcoma; carcinoma;
KW   wound granulation.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /note= "signal peptide"
FT   Protein         19..745
FT                   /note= "Tek extracellular domain"
FT   Protein         746..977
FT                   /note= "Fc portion"
XX
PN   WO200075323-A1.
XX
PD   14-DEC-2000.
XX
PF   07-JUN-2000; 2000WO-US15706.
XX
PR   07-JUN-1999;   99US-0137889.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Cerretti DP,  Borges LG,  Fanslow WC;
XX
DR   WPI; 2001-112149/12.
XX
PT   New Tek polypeptides antagonist having a fragment of the Tek
PT   extracellular domain, useful for treating mammals with a disease
PT   mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
PT   inflammatory diseases -
XX
PS   Claim 41; Page 34-37; 43pp; English.
XX
CC   The present sequence represents a fusion protein comprising the
CC   extracellular domain of the human receptor tyrosine kinase Tek and the
CC   Fc portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all
CC   or part of the region containing fibronectin type III (FNIII) motifs, and
CC   retains the ability to bind at least one Tek ligand. The fusion
CC   polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The
CC   polypeptide or soluble Tek multimer, antibodies or antibody fragments are
CC   useful for treating a mammal having a disease or condition mediated by
CC   angiogenesis, e.g. a solid tumour a condition or disease characterised by
CC   ocular neovascularisation. In particular, the Tek antagonists are useful
CC   for treating or preventing inflammatory diseases (e.g. arthritis,
CC   rheumatism or psoriasis), certain eye disorders (e.g. diabetic
CC   retinopathy, retinopathy of prematurity, neovascular glaucoma,
CC   retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
CC   degeneration or graft neovascularisation), cancer (e.g. metastatic
CC   sarcomas or carcinomas), or wound granulation.
XX
SQ   Sequence   977 AA;
```

```
  Query Match             61.4%;  Score 35;  DB 22;  Length 977;
  Best Local Similarity   60.0%;  Pred. No. 5.6e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATDIKGAEC 10
           |||  || :|
Db     289 catqwkqlqc 298
```

```
RESULT   5
AAR45440
ID   AAR45440 standard; Protein; 1124 AA.
XX
AC   AAR45440;
XX
DT   25-JUN-1994  (first entry)
XX
DE   Human orphan receptor kinase.
XX
KW   Ork; ligands; antibodies; PCR; amplification.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Peptide         1..18
FT                   /note  "signal peptide"
FT   Region          746..772
FT                   /note  "transmembrane region"
FT   Region          211..340
FT                   /note  "EGF-like repeat"
FT   Misc-difference 44
FT                   /note  "Cys of immunoglobulin domain"
FT   Misc-difference 102
FT                   /note  "Cys of immunoglobulin domain"
FT   Misc-difference 19..1124
FT                   /note= "claimed fragment"
FT   Misc-difference 19..745
FT                   /note  "claimed fragment"
XX
PN   WO9400469-A
XX
PD   06-JAN-1994
XX
PF   25-JUN-1993;    93WO-US06093.
XX
PR   26-JUN-1992;    92US-0905600.
XX
PA   (IMMV ) IMMUNEX CORP.
XX
PI   Ziegler SF;
XX
DR   WPI; 1994-026132/03.
DR   N-PSDB; AAQ55179.
XX
PT   DNA and protein sequences for orphan receptor tyrosine kinase -
PT   and expression vectors for prodn. of recombinant protein and
PT   antibodies specific for the protein, useful in research
XX
PS   Claim 13; Fig 1; 57pp; English.
XX
CC   Degenerate oligonucleotide primers based on the sequence conserved
CC   in the kinase domain of all receptor tyrosine kinases was used for
CC   PCR of single stranded cDNA from human placental polyA mRNA.  PCR
CC   prod. HTK 6 contained a novel sequence which was used as a probe to
CC   isolate longer fragments from a human placental cDNA library.  One
CC   clone contained the entire coding region and was called the ork gene.
CC   The gene prod. shown can be used as a research tool in in vitro assays
CC   for detection of ork, its ligands or their interactions.
CC   See also AAB45441.
XX
SQ   Sequence   1124 AA;


  Query Match             61.4%;  Score 35;  DB 15;  Length 1124;
  Best Local Similarity   60.0%;  Pred. No. 6.4e-02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAIDIKCAEC 10
           |||  || :|
Db     289 catqwkqlqc 298


RESULT   6
AAR73953
ID   AAR73953 standard; Protein; 1124 AA.
XX
AC   AAR73953;
XX
DT   22-JAN-1996  (first entry)
XX
DE   Human TEK tyrosine kinase protein.
XX
KW   tie 2; receptor-tyrosine kinase; DNA primer; cancer; angiogenesis;
KW   vasculogenesis; tek.
XX
OS   Homo sapiens.
XX
PN   WO9513387-A1.
XX
PD   18-MAY-1995.
XX
PF   12-NOV-1994;    94WO-EP03767.
XX
PR   12-NOV-1993;    93US-0152562.
XX
PA   (PLAC ) MAX PLANCK GES F FORDERUNG WISSENSCHAFTEN.
XX
PI   Risau W;
XX
DR   WPI; 1995-194125/25.
XX
PT   New tie 2 receptor tyrosine kinase and related nucleic acid - and
PT   methods for detecting tie 2 modulators for treating or cancer,
PT   associated with angiogenesis and vasculogenesis
XX
PS   Disclosure; Page 49; 81pp; English.
XX
CC   This protein is the human homolog of mouse tie 2 receptor
CC   tyrosine kinase.
XX
SQ   Sequence   1124 AA;


  Query Match             61.4%;  Score 35;  DB 15;  Length 1124;
  Best Local Similarity   60.0%;  Pred. No. 6.4e-02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAIDIKCAEC 10
           |||  || :|
Db     289 catqwkqlqc 298


RESULT   7
AAY80318
ID   AAY80318 standard; protein; 1124 AA.
XX
AC   AAY80318;
XX
DT   15-NOV-1999  (first entry)
XX
DE   Amino acid sequence of TEK (also known as TIE.) protein.
XX
KW   TEK protein; TIE2 protein; receptor tyrosine kinase; T-cell receptor;
KW   immune response; endothelial cell; tumor associated vasculature;
KW   coagulation; thrombosis; cancer; anticancer therapy.
XX
OS   Homo sapiens.
XX
PN   WO9943801-A1.
XX
PD   02-SEP-1999.
XX
PF   26-FEB-1999;    99WO-GB000584.
XX
PR   26-FEB-1998;    98GB-0004121.
XX
PA   (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PI   Durrant LG,  Hewett PW,  Ramage JM,  Spendlove I.
XX
DR   WPI; 1999-540586/45.
XX
PT   New peptides containing at least one epitope from TEK receptor
PT   tyrosine kinase, used in vaccines against cancer
XX
PS   Disclosure; Fig 1; 54pp; English.
XX
CC   The present sequence represents the TEK protein, one of a family of
CC   protein. TEK is a receptor tyrosine kinase. TEK is found on cell pro-
```

CC which bind to MHC. The presentation of the epitopes can also stimulate
CC helper cell and/or cytotoxic T cell responses. The immune response is
CC directed against endothelial cells in the tumor-associated
CC vasculature and includes production of antibodies that bind to the
CC cells, causing coagulation and thrombosis. The immune response is
CC targeted to endothelial cells lining blood vessels of the tumor (these
CC cells overexpress Tek). So damage to even a few cells will kill many
CC tumor cells. These target cells are accessible to the immune response
CC and problems of antigenic heterogeneity, MHC loss and resistance to
CC apoptosis (associated with epithelial cells) are unlikely to occur in
CC normal endothelial cells. TEK epitopes (see AAY10329.24) are used to
CC generate antibodies, and for prevention and treatment of cancer.
CC The peptides, and recombinant DNA constructs or viral vectors that
CC express them, are useful as anticancer vaccines to target endothelial
CC cells that line blood vessels of the tumor.
XX
SQ Sequence 1124 AA;

Query Match 61.4%; Score 35; DB 20; Length 1124;
Best Local Similarity 60.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATDIKGAEC 10
|||.|| :|
Db 289 catqwkqlpc 298

RESULT 8
AAW88448
ID AAW88448 standard; Protein; 1296 AA.
XX
AC AAW88448;
XX
DT 26-APR-1999 (first entry)
XX
DE Caenorhabditis elegans NPC1 protein orthologue.
XX
KW Niemann-Pick disease type C; NPC1 gene; worm; orthologue;
KW diagnosis, therapy, animal model, cholesterol, neurodegeneration.
XX
OS Caenorhabditis elegans.
XX
PN WO9901555-A1.
XX
PD 14-JAN-1999.
XX
PF 02-JUL-1998; 98WO-US13802.
XX
PR 03-JUL-1997; 97US-0051682.
XX
PA (USSH) US DEPT HEALTH & HUMAN RESOURCES.
XX
PI Carstea ED, Gu J, Loftus SK, Morris JA, Pavan WJ;
PI Pentchev PG, Rosenfeld MA, Tagle DA;
XX
DR WPI; 1999-106056/09.
DR N-PSDB; AAX06876-77.
XX
PT New isolated gene, NPC 1 is associated with Niemann Pick type C
PT disease, used to develop products for the study, diagnosis and
PT therapy of the disease
XX
PS Disclosure; Page 91-95; 101pp; English.
XX
CC This polypeptide comprises the Caenorhabditis elegans orthologue
CC of the human NPC1 polypeptide that is associated with
CC Niemann Pick disease type 2 (NP C). The polypeptide shows
CC extensive identity (30%) and similarity (55%) to the human
CC NPC1 protein (see AAW88445). Biochemical and genetic analysis of
CC yeast, worm and murine NPC1 model systems will provide resources
CC for understanding the role of NPC1 in intracellular cholesterol
CC homeostasis and in the aetiology of neurodegeneration in NP C
CC disease. The provision of the human NPC1 DNA sequence (see
CC AAX06873) enables methods of detecting the presence of mutations in
CC the hNPC1 gene, and thereby facilitates the determination of
CC whether an individual is an NP-C sufferer or carrier.
XX
SQ Sequence 1296 AA;

Query Match 60.7%; Score 34.5; DB 20; Length 1296;
Best Local Similarity 58.8%; Pred. No. 8.9e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

Qy 1 CAT---DIKGAEC 10
|| | |:|| |
Db 1006 cvtlqiidvkqaac 1018

RESULT 9
AAR83764
ID AAR83764 standard; peptide; 9 AA.
XX
AC AAR83764;
XX
DT 18-MAR-1996 (first entry)
XX
DE NGF derived structural analog 1(28-36) binds neurotrophin receptor.
XX
KW Neurotrophin receptor; structural analog; nerve growth factor; beta turn;
KW reverse turn; cyclic; tracer, uncharged; hydropathic; inhibition;
KW neurite outgrowth; central nervous system; peripheral nervous system;
KW tumour; neuroma; hormone-receptor interaction site; immunisation;
KW receptor domain-function correlation.
XX
OS Synthetic.
XX
PN WO9521198-A1.
XX
PD 10-AUG-1995.
XX
PF 07-FEB-1995; 95WO-CA00059.
XX
PR 07-FEB-1994; 94GB-0002331.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Cuello AC, Lesauteur L, Saragovi UH;
XX
DR WPI; 1995-283721/37.
XX
PT New cyclic peptide(s) which bind to neurotrophin receptor and mimic
PT or inhibit neurotrophin activity useful e.g. for inhibiting
PT neurite outgrowth or treating nervous system disease, tumours, etc.
XX
PS Claim 9; Page 32; 42pp; English.
XX
CC The peptides AAR83760-80 are examples of peptides that bind to the
CC neurotrophin receptor under physiological conditions in vivo or in
CC vitro. The peptides are structural analogs of nerve growth factor (NGF)
CC and contain at least one beta turn (from region 28-31, 43-46 or 91-94)
CC or 3 consecutive reverse turns (from region 59-65). The peptides may be
CC cyclised by amide or S-S or other cyclisation procedures and may be
CC linked to a tracer e.g. a metal chelate or radionuclide. The peptides
CC may contain uncharged or hydropathic amino acids at the N- and
CC C-termini, which are optionally protected by an FMOC or acetyl protecting
CC group. The peptides are useful for inhibition of neurotrophin, esp. for
CC inhibiting neurite outgrowth, treating central or peripheral nervous
CC system disease, tumours and neuromas, for mapping hormone-receptor
CC interaction sites and receptor domain-function correlation and for
CC immunisation. This peptide is derived from residues 28-36 and is
CC linear.
XX
SQ Sequence 9 AA;

Query Match 59.6%; Score 84; DB 16; Length 9;
Best Local Similarity 87.5%; Pred. No. 3.4e-05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AIDIKGAE 9
|||||||:|
Db 1 atdikqke 8

RESULT 10
AAR83761
ID AAR83761 standard; peptide; 10 AA.
XX
AC AAR83761;
XX
DT 18-MAR-1996 (first entry)
XX
DE NGF derived structural analog C(30-35) binds neurotrophin receptor.
XX
KW Neurotrophin receptor; structural analog; nerve growth factor; beta turn;
KW reverse turn; cyclic; tracer; uncharged; hydropathic; inhibition;
KW neurite outgrowth; central nervous system; peripheral nervous system;
KW tumour; neuroma; hormone-receptor interaction site; immunisation;
KW receptor domain-function correlation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note "any uncharged or hydropathic amino acid"
FT Modified-site 1
FT /note "contains Fmoc protective group"
FT Misc-difference 10
FT /note "any uncharged or hydropathic amino acid"
XX
PN WO9521194-A1.
XX
PD 10-AUG-1995.
XX
PF 07-FEB-1995; 95WO-CA00059.
XX
PR 07-FEB-1994; 94GB-0002331.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Cuello AC, Lesauteur L, Saragovi UH;
XX
DR WPI; 1995-283741/37.
XX
PT New cyclic peptide(s) which bind to neurotrophin receptor and mimic
PT or inhibit neurotrophin activity - useful e.g. for inhibiting
PT neurite outgrowth or treating nervous system disease, tumours, etc.
XX
PS Claim 9; Page 82; 42pp; English.
XX
CC The peptides AAR83760-80 are examples of peptides that bind to the
CC neurotrophin receptor under physiological conditions in vivo or in
CC vitro. The peptides are structural analogs of nerve growth factor (NGF)
CC and contain at least one beta turn (from region 28-36, 43-49 or 91-98)
CC or 3 consecutive reverse turns (from region 59-65). The peptides may be
CC cyclised by oxidn. of Cys or other cyclisation procedure and may be
CC linked to a tracer e.g. a metal chelate or radionuclide. The peptides
CC may contain uncharged or hydropathic amino acids at the N- and
CC C-termini, which are optionally protected by an Fmoc or acetyl protecting
CC group. The peptides are useful for inhibition of neurotrophin, esp. for
CC inhibiting neurite outgrowth, treating central or peripheral nervous
CC system disease, tumours and neuromas, for mapping hormone-receptor
CC interaction sites and receptor domain-function correlation and for
CC immunisation. This peptide is derived from residues 30-35 and contains
CC a beta turn.
XX
SQ Sequence 10 AA;

Query Match 59.6%; Score 84; DB 16; Length 9;
Best Local Similarity 87.5%; Pred. No. 3.4e-05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DIKGAEC 10
|||| ||
Db 3 dikqkec 9

RESULT 11
AAP90623
ID AAP90623 standard; peptide; 11 AA.
XX
AC AAP90623;
XX
DT 10-FEB-1998 (first entry)
XX
DE Peptide to induce nerve growth factor response.
XX
KW Rat dorsal root ganglia; Alzheimer's disease; Lowe's syndrome; Rett
KW syndrome; Creutzfeldt-Jacob disease; Gerstmann-Straussler syndrome;
KW Huntington's chorea; neuroblastoma.
XX
OS Synthetic.
XX
PN EP835657-A.
XX
PD 04-APR-1998.
XX
PF 28-MAR-1989; 98EP-0302056.
XX
PR 28-MAR-1988; 88US-0000698.
XX
PA (REGE-) UNIV CALIFORNIA.
XX
PI Longo FM, Mobley WC, Kaiser SC;
XX
DR WPI; 1989-287098/40.
XX
PT New nerve growth factor peptide(s)
PT - with agonist and/or antagonist activity
XX
PS Disclosure; Table 1; page 7; 34pp; English.
XX
CC This corresponds to amino acids 20-48 (esp. 4-11), a human nerve growth
CC factor, and esp. has the terminal alpha-CO2H grp. amidated. Has nerve
CC (anti)agonist activity for treating neurodegenerative diseases.
XX
SQ Sequence 11 AA;

Query Match 54.6%; Score 84; DB 16; Length 11;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AIDIKGAE 9
|||||||:|
Db 1 atdikqke 8

RESULT 12
AAR41747
ID AAR41747 standard; peptide; 12 AA.
XX
AC AAR41747;
XX
DT 18-MAR-1994 (first entry)
XX
DE Conserved NGF region 29-36.
XX
KW Nerve growth factor; NGF; brain-derived neurotrophic factor; BDNF;

KW neurotrophin; NT; p75NGFR; receptor; transmembrane glycoprotein;
KW rat; mouse; human; bovine; guinea pig; chicken;
KW xenopus; snake; mutant; modification.
XX
PN WO9318066-A.
XX
PD 16 SEP 1993.
XX
PF 08 MAR 1993; 93WO-SE00201.
XX
PR 06-MAR-1992; 92US-0847369.
XX
PA (MOLI/) MOLINER C F I.
PA (PERS/) PERSSON H B.
XX
PI Persson HB, Moliner CFI;
XX
DR WPI; 1993-303406/38.
XX
PT New mutant forms of neurotrophic factor of increased stability
PT have specific positively charged aminoacid(s) replaced, showing
PT more selective binding but equiv. biological activity
XX
PS Disclosure; Fig 1; 70pp; English.
XX
CC Comparison of NGF genes from different species has revealed clusters
CC of amino acid residues which are highly conserved across different
CC gps. of vertebrates (rat, mouse, human, bovine, guinea pig and chicken
CC NGF: AAR41747; xenopus NGF: AAR41748; snake NGF: AAR41749; rat BDNF:
CC AAR41750; rat NT 3: AAR41751; and xenopus NT-4: AAR41752).
CC New mutant neurotrophic factor (NF) comprises a wild-type NF having,
CC as modification, replacement of one or more positively charged amino
CC acids, in the amino acid regions 30-34 or 93-98, with uncharged or
CC negatively charged amino acids, so that the mutants have reduced
CC ability to bind to p75NGFR (a transmembrane glycoprotein which binds
CC NGF with low affinity); c.f. wild type NF.
CC The wild-type NF is pref. NGF, BDNF, NT 3 or NT 4. In the 30-34
CC region, the modifications are pref. at Lys32 (or Arg32 in NT)
CC and/or Lys34, in the 93-98 region, Lys95 is partic. replaced by Ala.
CC Esp. for NT 3, Arg32, His34, Asn93 and/or Asn94, for BDNF, Lys95,
CC Lys96 and/or Arg97, and for NT 4 Glu94 and Arg96 are replaced.
CC The mutants differ from wild-type NF as regards receptor binding
CC affinity and specificity, esp. they can bind to trk receptor but not
CC to p75NGFR. Modifications in the 25-36 amino acid region improve
CC stability.
XX
SQ Sequence 12 AA;

Query Match 59.6%; Score 34; DB 14; Length 12;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAE 9
|||||| |
Db 4 atdikgke 11

RESULT 13
AAR41748
ID AAR41748 standard; peptide; 12 AA.
XX
AC AAR41748;
XX
DT 18-MAR-1994 (first entry)
XX
DE Conserved NGF region 25-36 (xenopus).
XX
KW Nerve growth factor; NGF; brain-derived neurotrophic factor; BDNF;
KW neurotrophin; NT; p75NGFR; receptor; transmembrane glycoprotein;
KW rat; mouse; human; bovine; guinea pig; chicken;
KW xenopus; snake; mutant; modification.
XX
PN WO9318066-A.
XX
PD 16 SEP 1993.
XX
PF 08-MAR-1993; 93WO-SE00201.
XX
PR 06-MAR-1992; 92US-0847369.
XX
PA (MOLI/) MOLINER C F I.
PA (PERS/) PERSSON H B.
XX
PI Persson HB, Moliner CFI;
XX
DR WPI; 1993-303406/38.
XX
PT New mutant forms of neurotrophic factor of increased stability -
PT have specific positively charged aminoacid(s) replaced, showing
PT more selective binding but equiv. biological activity
XX
PS Disclosure; Fig 1; 70pp; English.
XX
CC Comparison of NGF genes from different species has revealed clusters
CC of amino acid residues which are highly conserved across different
CC gps. of vertebrates (rat, mouse, human, bovine, guinea pig and chicken
CC NGF: AAR41747; xenopus NGF: AAR41748; snake NGF: AAR41749; rat BDNF:
CC AAR41750; rat NT 3: AAR41751; and xenopus NT-4: AAR41752).
CC New mutant neurotrophic factor (NF) comprises a wild type NF having,
CC as modification, replacement of one or more positively charged amino
CC acids, in the amino acid regions 30-34 or 93-98, with uncharged or
CC negatively charged amino acids, so that the mutants have reduced
CC ability to bind to p75NGFR (a transmembrane glycoprotein which binds
CC NGF with low affinity), c.f. wild type NF.
CC The wild-type NF is pref. NGF, BDNF, NT 3 or NT 4. In the 30-34
CC region, the modifications are pref. at Lys32 (or Arg32 in NT)
CC and/or Lys34, in the 93-98 region, Lys95 is partic. replaced by Ala.
CC Esp. for NT 3, Arg32, His34, Asn93 and/or Asn94, for BDNF, Lys95,
CC Lys96 and/or Arg97, and for NT 4 Glu94 and Arg96 are replaced.
CC The mutants differ from wild-type NF as regards receptor binding
CC affinity and specificity, esp. they can bind to trk receptor but not
CC to p75NGFR. Modifications in the 25-36 amino acid region improve
CC stability.
XX
SQ Sequence 12 AA;

Query Match 59.6%; Score 34; DB 14; Length 12;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAE 9
|||||| |
Db 4 atdikgke 11

RESULT 14
AAR07158
ID AAR07158 standard; protein; 13 AA.
XX
AC AAR07158;
XX
DT 24-JAN-1991 (first entry)
XX
DE Nerve growth factor (NGF) peptide fragment.
XX
KW Nervous disorders; Alzheimer's disease; Parkinson's disease; stroke.
XX
OS Homo sapiens.
XX
PN WO9010644-A.
XX
PD 20-SEP-1990.
XX
PF 08-MAR-1990; 90WO-0000149.

XX
PR 14-MAR-1989; 89SE-0000899.
XX
PA (LOPE-) LOPE MED AB.
XX
PI Olson L, Persson H, Ebendal T;
XX
DR WPI; 1990-304983/40
XX
PT New peptide fragments of nerve growth factor or its precursor
PT used to raise specific antibodies for immunoassay, esp. for brain
PT tissue
XX
PS Claim 1; Page 13; 24pp; English.
XX
CC Peptides are Abs raised to them are useful in detecting the presence
CC of NGF and precursors, allowing early diagnosis and treatment of
CC nervous disorders eg. Alzheimer's and Parkinson's disease, spinal
CC cord injury, stroke etc.
CC Peptide corresponds to AAs 23 to 35 of human NGF mature protein.
XX
SQ Sequence 13 AA;

Query Match 59.6%; Score 34; DB 11; Length 13;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAE 9
|||||| |
Db 6 atdikgke 13

RESULT 15
AAR07161
ID AAR07161 standard; protein; 14 AA.
XX
AC AAR07161;
XX
DT 24 JAN-1991 (first entry)
XX
DE Synthetic Nerve growth factor (NGF) peptide fragment.
XX
KW Nervous disorders, Alzheimer's disease, Parkinson's disease, stroke.
XX
OS Synthetic.
XX
PN WO9010644-A.
XX
PD 20 SEP-1990.
XX
PF 08 MAR-1990; 90WO-0000149.
XX
PR 14 MAR-1989; 89SE-0000899.
XX
PA (LOPE-) LOPE MED AB.
XX
PI Olson L, Persson H, Ebendal T;
XX
DR WPI; 1990-304983/40.
XX
PT New peptide fragments of nerve growth factor or its precursor -
PT used to raise specific antibodies for immunoassay, esp. for brain
PT tissue
XX
PS Disclosure; Page 11; 24pp; English.
XX
CC Peptides are Abs raised to them are useful in detecting the presence
CC of NGF and precursors, allowing early diagnosis and treatment of
CC nervous disorders eg. Alzheimer's and Parkinson's disease, spinal
CC cord injury, stroke etc.
CC Peptide corresponds to AAs 23 to 35 of rat (mouse, chicken, human) NGF.
XX
SQ Sequence 14 AA;

Query Match 59.6%; Score 34; DB 11; Length 14;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ATDIKGAE 9
|||||| |
Db 7 atdikgke 14

Search completed: August 30, 2001, 16:41:24
Job time: 27 sec